A Bayesian Space-Time Dynamic Linear Model for Radioactivity Deposition after a Nuclear Accident

Thesis

A Bayesian Space-Time Dynamic Linear Model for Radioactivity Deposition after a Nuclear Accident

by

Swarup De

B.Sc. Hons. (Statistics)
M.Sc. (Statistics)

A dissertation submitted in partial satisfaction
of the requirements for the degree of

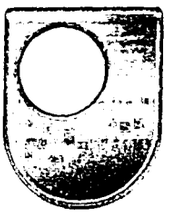
Doctor of Philosophy

in

Statistics

Department of Mathematics and Statistics
The Open University
Milton Keynes
United Kingdom

August 10, 2009



The Open University

EX12

RESEARCH SCHOOL

Library Authorisation Form

Please return this form to the Research School with the two bound copies of your thesis to be deposited with the University Library. All candidates should complete parts one and two of the form. Part three only applies to PhD candidates.

Part One: Candidates Details

Name: *SWARUP DE* PI: ***X5411467***

Degree: ***Doctor of Philosophy***

Thesis title: ***A Bayesian Space-Time Dynamic Linear Model for Radioactivity Deposition after a Nuclear Accident***

Part Two: Open University Library Authorisation

I confirm that I am willing for my thesis to be made available to readers by The Open University Library, and that it may be photocopied, subject to the discretion of the Librarian.

Signed: Swarup De Date: 13 August 2009

Part Three: British Library Authorisation [PhD candidates only]

If you want a copy of your PhD thesis to be available on loan to the British Library Thesis Service as and when it is requested, you must sign a British Library Doctoral Thesis Agreement Form. Please return it to the Research School with this form. The British Library will publicise the details of your thesis and may request a copy on loan from the University Library. Information on the presentation of the thesis is given in the Agreement Form.

Please note the British Library have requested that theses should be printed on one side only to enable them to produce a clear microfilm. The Open University Library sends a soft bound copy of theses to the British Library.

The University has agreed that your participation in the British Library Thesis Service should be voluntary. Please tick either (a) or (b) to indicate your intentions.

(a) ☑ I am willing for The Open University to loan the British Library a copy of my thesis. A signed Agreement Form is attached

(b) ☐ I do not wish The Open University to loan the British Library a copy of my thesis.

Signed: Swarup De Date: 13 August 2009

ABSTRACT

A three-stage Hierarchical Bayesian Space-Time (HBST) model, an extension of the class of dynamic linear models to space, is proposed for the ground contamination levels (GCL) and related uncertainties caused by polluting discharges in the environment. The model should allow updating the distribution of GCL in time and space as new data in the form of measurements and expert judgements become available to give real-time estimates of deposition levels.

An application of the HBST model is proposed for the statistical modelling of radioactivity deposition after a nuclear accident. It explicitly handles uncertainties associated with (i) predictions of depositions from a long-range atmospheric dispersal model, (ii) in-situ gamma ray measurements and (iii) spatial interpolations. Unlike existing environmental statistical models, the HBST model also accounts for an established food chain contamination model called ECOSYS for which it provides data assimilation capabilities.

The HBST model permits a fast implementation and full probabilistic inference for the parameters, interpolation and forecasts. Three distinct formulations of the HBST model were applied to assimilate real data of radioactivity deposition from the Chernobyl accident in southern Germany. Two of those formulations differ on the functional form of their spatial covariance matrices while the third, a normal inverse-Wishart model, allows the spatial covariances to "learn" from

the data within the usual Bayesian paradigm. The later is shown to outperform the former models both in short and medium term forecasting as well as in a predictive interpolation test that took some measurements as out-of-sample.

DEDICATION

This research study is dedicated to the memory of my brother Anup De

and

to my parents & other family members

for letting me pursue my dream

for so long

so far away from home

ACKNOWLEDGEMENTS

I express sincere appreciation to Dr Alvaro E Faria and Prof Kevin J McConway, for supporting me over the years, and for giving me encouragement, patience and positive criticism. I would like to thank Dr Nickolay T Trendafilov and Dr Catriona Queen for their support. Thanks go to the other faculty members for their valuable suggestions and comments, and for providing an excellent atmosphere for research.

Weekly meetings with supervisor(s) have proved to be one of my best learning experiences at The Open University and have kept me honest about all the details. Financial support from The Open University is gratefully acknowledged.

I would like to thank the Bavarian Regional Office for Environmental Protection to provide the real radioactivity data of the 1986 Chernobyl's accident contamination of Bavaria in southern Germany.

I express my thanks and appreciation to all the members in the computer support team in our faculty. In particular, Mr Robin Goodman and Mrs Maja Dunn helped me quite a lot over the years. I would like to thank our former departmental secretary Mrs Val Spearman, current secretary Mrs Sarah Frain, Mrs Tracy Johns and Miss Sara Griffin for their support over the years. I would also like to thank Mr Thomas Ferrere from Academic & Administrative Computing Services (AACS) for his support.

I would like to thank my friends and colleagues at The Open University

with whom I have had the pleasure of working over the years. In particular, I would like to express my gratitude to Emmanuel Mubwandarikwa, Lin Sun, David Jenkinson, Steffen Unkel, Doyo Gragn, Angela Noufaily and Fadlalla Elfadaly. I would like to thank few of my best friends like Olivier Zanellato and Evelyn Wandia who always supported and encouraged me over the years. Also, I would like to thank my friends in India and beyond for providing a welcome distraction from college.

Finally and most importantly, I express my thanks and appreciation to my parents, my brothers & sisters and Anasuya for giving me the motivation to finish this thesis. They are the closest to my heart. In all the ups & down, they always supported me. In the second year of my PhD, I lost my brother who was my best friend. He always motivated me and encouraged me. I would like to dedicate my PhD to him.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Setting the Scene

This study proposes and tests a statistical model for uncertainty handling and data assimilation associated with the ground deposition of radioactivity after a nuclear accident. The model makes use of all information available during the accident in order to provide the best possible estimation of the levels of ground contamination by the released radioisotopes (together with a measure of their quality) for decision support purposes at various levels in nuclear emergencies, such as distribution of iodine tablets, sheltering advice and evacuation. For this, a hierarchical Bayesian spatio-temporal (HBST) model was developed to coherently combine information coming from established physical and environmental models with in-situ near ground radioactivity measurements.

The scenario considered here is one of a severe accident in a nuclear plant with radioactive releases into the environment. This can typically be characterized as having four distinct phases:

(1) a pre-release phase, when the in-plant conditions deteriorate in such a way

that the accidental release of radioactive material (source term) is about to occur;

(2) a release phase, when while being released the source term is dispersed into the atmosphere;

(3) an early post-release phase, when while being dispersed, the released masses of radionuclides are also partially deposited in time at diverse locations, contaminating soil, plants, waters, and thus, entering the human food chain; and

(4) a late post-release phase where contamination passes through animal and food chains and impacts on long term environmental factors.

The pre-release and release phases (1) and (2) above are associated with the early stages of the accident, that is, with a time scale ranging from few minutes to one or two days of duration. The post-release phases (3) and (4) are associated with the medium and long term and time scales ranging from days to years.

Diverse deterministic but typically very complex physical and environmental models have been developed to describe a nuclear accident with its related contamination and effects during each phase. Examples of those for the early phase are the in-plant status probabilistic belief net for source term estimation (Smedley *et al.*, 1996) and the Lagrangian K-model model for long-range atmospheric dispersal (Lauritzen *et al.*, 2006). For the later stages, the ECOSYS (Müller and Pröhl, 1993) describes the food chain contamination. Because of

their hazardous environmental and health effects, any auxiliary experiments that have been performed are either limited in their association with the real scenario (for example, tracer experiments which use non-toxic smoke emissions typically in constant wind field environment) or are the product of strict control (like lab tests on plants exposure to radiation). The result of this is that current physical and environmental models are largely descriptive.

Data from the rare accidents that have occurred are not generally available for two reasons, the sensitivity of such information and the lack of proper measurement protocols during the emission (it is only in the last few years that these protocols have been implemented across Europe for example). When available they are useful for the modelling related to the geographic region they come from but not for other regions.

Typically, such models have great complexity, lack of linearity and an internal integrity. They can be thought as machines which take a function of outputs from previous modules (except for the first), vectors of observations and vectors of known covariates as their inputs to produce outputs. The outputs of this module can then be used as one of the components of the inputs of a subsequent module.

Part of the work we describe here has been developed in the RODOS (Real time Online Decision Support) system, an EU project for decision support in nuclear emergencies, that includes a module for each of the above phases.

Similarly to Gamerman *et al.* (2006), the HBST model adopt a state-space approach based on an extension of the dynamic linear model of West and Har-

rison (1997) to our spatio-temporal process. However, unlike Gamerman *et al.* (2006), the proposed approach in this thesis includes specific environmental and physical models in its formulation. In fact, the development of the integrated hierarchical Bayesian approach describes in this thesis is oriented to include the ECOSYS model of Müller and Pröhl (1993) as a particular case. The ECOSYS is a detailed physical-biological model developed to explain the spread of radioactive contamination within the food chain with possible consequences on the human exposure doses. It includes modules for ground deposition, food chain contamination and human exposure. The ground deposition module, the one I will be focusing on in this thesis, assesses the radiological consequences of short-term depositions of radio-nuclides in diverse types of soil and plants. The proposed HBST model handles in a rather natural and structured way the uncertainties associated with the ECOSYS model as we shall see. In addition, this model allows fairly easily not only the combination of information coming from different sources (such as estimations based on atmospheric dispersal and near-ground radioactivity measurements) but also the inclusion of new components, when needed (such as trends and seasonality for example), by augmentation of vectors and matrices only. Also, at any time the probability density of all variables will be dynamically updated by the information available up to that time. That is, the joint density is updated as new data become available with no need for the full data set to be collected before updating can take place. Moreover, different types of measured data need not all be available at every monitoring station as usually will be the

case. In practice, different stations will have different measuring resources, for example, stations equipped to measure γ-dose rates not necessarily will be prepared to measure air concentration and/or amount of rainfall. To handle nuclide specific wet and dry deposited activities on soil, the HBST model was formulated with three main components. Those components are represented by variables and parameters associated with total deposited activity and proportions of both specific radioisotopes and wet deposition relative to the total ground deposition mass at each site of interest.

Because the HBST model allows learning (from data) on the covariance structure related to the model, part of the results in this thesis are comparable to those of Le and Zidek (1992). Basically, it consists of the application of standard multivariate normal-wishart conjugate analysis (Anderson, 1984). However, differently from Le and Zidek (1992), in the case here significant temporal structures are included (temporal autocorrelation is allowed) as well as an explicit direct modelling of spatial interpolations. The latter allows for smoothness constraints on the covariance estimates through the specification of a suitable spatial interpolation matrix.

The HBST model focuses on the interface between components of the atmospheric dispersal, K-model, and the food chain, ECOSYS. We shall make use of outputs from the K-model, such as its predictions of radioactivity deposition for a regular spatial grid, as inputs to the HBST model. Those predictions are treated as data and combined with in-situ deposition measurements taken at

sites which are usually off-grid locations. In practice, the deposition measurements (or other measurements from which ground deposition can be calculated, such as near-ground air concentrations together with rainfall intensities and near-ground gamma dose rates) will not be available for all sites of interest. In fact, there will be a number of fixed measuring stations irregularly and sparsely separated over the region of interest. Some mobile measuring capability should be possible however in small scale. Also, there is a subtle interface question between the atmospheric dispersal K-model and the ECOSYS which is that of what deposition is meant in each model. Because, deposition for ECOSYS is not only nuclide but also plant specific, I assume (for the models' reconciliation sake) that near-ground deposition at a certain site corresponds to deposition on soil.

The proposed HBST model was applied to real data of the 1986 Chernobyl's accident contamination of Bavaria in southern Germany. In fact, three versions of the HBST model combined predictions at fixed regular grid points from the long range atmospheric K-model (Lauritzen and Mikkelsen, 1999) mentioned in Chapter 6 with measured (off-grid) near-ground contamination in Bavaria to produce estimations (actual and predictive) of the contamination profile.

The accident at the Chernobyl nuclear reactor (in northern Ukraine, near the border with Belarus) that occurred on 26 April 1986 was the most serious accident ever to occur in the nuclear power industry. The accident happened during an experimental test of the electrical control system when the reactor was being shut down for routine maintenance. A sudden power surge caused a chain

of events leading to the destruction of the reactor's core, the severe damaging of the plant's building and ultimately to the release of considerable amounts of radioactive materials to the atmosphere. The leaking core was exposed to the elements for ten days before engineers were able to seal the building. The radioactive gases and particles released in the accident were initially carried by the wind in westerly and northerly directions. On subsequent days, distinct wind directions prevailed. The deposition of radionuclides was primarily governed by any precipitation occurring during the passage of the radioactive cloud, leading to a complex and variable exposure pattern throughout the affected region and as well as to other European countries including Germany as we shall see in the later chapters.

The following subsections introduce a brief history of the spatial modelling with the basic introduction of some keywords which have been used in the later chapters.

1.2 A Brief History of the Spatial Modelling

In the history of spatial modelling, Student (1907) first gives the concept of spatial modeling. He was concerned about the distribution of particles throughout a liquid. Instead of analysing the spatial position of the particles, he was more interested to find the pattern of the particles in the liquid. In this context, he aggregated the data into counts of particles of per unit area. A hemocytometer (which is a device to count different types of cells as well as other microscopic

particles) of area 1 mm^2, divided into 400 squares, was used to count yeast cells. He found that the distribution of the number of cells per square followed a Poisson distribution.

After Student's experiment (Student, 1907), it was clear that the spatial data have probability distributions. In other words models exists for spatial data. But then the question arose of whether any dependency exists between the spatial data. R. A. Fisher was clearly aware of spatial dependence in agricultural field experiments, since he tried to removed it (Fisher, 1935). In that period, he established the principles of randomization, blocking and replication. Randomization controls the unwanted bias and also neutralizes (but does not remove) the effect of spatial correlation (Yates, 1938). However, Fisher realized that randomization does not neutralize the spatial correlation at spatial scales larger or smaller than the plot dimensions. After Fisher, Fairfield (1938) also empirically shows the presence of spatial correlation in agricultural field experiments.

Now in spatial modeling, the crucial thing is to include the spatial dependency into the modeling. Whittle (1954), first introduced the spatial dependency. According to Whittle (1954), the sampling theory of stationary processes (discussed in Section 1.2.3) in *space* is not completely analogous to that of stationary time series, since the variate of a time series is influenced only by past values, while for spatial process dependence extends in all directions.

Then the nearest-neighbor (NN) methods have been developed. Papadakis (1937) was the first to introduce a NN method for the analysis of agricultural field

trials. But the theoretical properties of NN methods were first investigated by Atkinson (1969). Nearest-neighbor methods for analysing agricultural field trials attempt to take spatial dependence into account, indirectly, by using residuals from neighboring plots as covariates (see Bartlett (1938), Bartlett (1978), Wilkinson *et al.* (1983), Green *et al.* (1985), Besag and Kempton (1986), Zimmerman and Harville (1989)).

But in some areas, such as geology, ecology and environmental science, it is not often possible (nor always appropriate) to randomize, block, and replicate the data. So, it needs new statistical models and approaches which address the new questions arising from old and new technologies.

In the following subsections, after a short introduction of spatial data, some terms are defines that are related to the spatial modeling.

1.2.1 A Short Introduction to Spatial Data

In this section, we will see the basic definition of different *spatial data.* According to Cressie (1993), we classify spatial data into three basic types. (i) *Geostatistical data*, (ii) *Lattice data* and (iii) *Point pattern data.*

The above mentioned types of data can be defined in the following way. Define $\mathbb{R}^d$ as the d-dimensional Euclidean space throughout the thesis. Let $\mathbf{s} \in \mathbb{R}^d$ be spatial location of the data and $\mathbf{Y}(\mathbf{s})$ be random variable for the data at the spatial location $\mathbf{s}$. Define, D as the index set and allow $\mathbf{s}$ to vary over $D \subset \mathbb{R}^d$.

So, now we can define the multivariate random field as

$$\{\mathbf{Y}(\mathbf{s}) : \mathbf{s} \in D\}$$

For *Geostatistical data*, D is a fixed subset of $\mathbb{R}^d$ that contains a d-dimensional rectangle of positive volume, i.e. continuous $\mathbf{s}$ and $\mathbf{Y}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$. E.g. scallops data. (The scallops data frame is a spatial data set listing the catch of scallops from a 1990 National Marine Fisheries Service trawl survey in the Atlantic Ocean. The survey area runs from the Delmarva Peninsula off the coast of Virginia and Maryland up to the George Banks.)

For *lattice data*, D is a fixed (regular or irregular) collection of countably many points of $\mathbb{R}^d$ and $\mathbf{Y}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$. E.g. image analysis where the image is recorded as pixels on a lattice.

And for *point patterns data*, D is (random) a point process in $\mathbb{R}^d$ or a subset of $\mathbb{R}^d$ and $\mathbf{Y}(\mathbf{s})$ is a random vector at location $\mathbf{s} \in D$. E.g. data on trees located randomly in forest.

The proposed model in Chapter 4 is based on lattice data, so we concentrate only on spatial lattice data.

Spatial lattice data is a set of quantitative measurements recorded on a regular lattice, for example measurements of ground contamination level by radioactivity after a nuclear accident at a given area taken at regular spacing in a geographical region of interest. In general, the lattice patterns arise in planned agricultural trials and in satellite imaging. In any spatial data set whose spatial locations

are form a regular lattice in the Euclidean space is the closest analogue to a time series observed at equally spaced time points. From the spatial lattice data, neighborhood information (based on, e.g., Euclidean distance) can be specified. The process that generates a lattice pattern is called a lattice process. Analysis of a lattice process is provided via either a spatial domain approach using *spatial auto-correlation* or a frequency domain approach using *spectral analysis*. Spatial auto-correlation methods measure correlation between measurements at pairs of sites (For further details please see, Upton and Fingleton (1995); Cliff and Ord (1981)).

In any statistical analysis of data in a spatial lattice, it should be determined whether the locations (i) are *regular* or *irregular*, (ii) represent *points* or *regions*, and (iii) are indices for *continuous* or *discrete* random variables. By nature or by design, many spatial problems are *regular*, *points* and *discrete*. The class of models in this thesis were formulated for *points* data at *regular* grids and *discrete* space (and / or *continuous*) in nature as we shall see. An example of spatial lattice data is given in Figure 1.1. Figure 1.1 shows a 64×64 regular grid covering the Bavarian region in southern Germany that was used by the proposed model in the application in Chapter 6. The distance between consecutive grid points is 8 Km and each grid point has its own data of radioactivity deposition which gives the map in Figure 1.1. Detail explanation of those maps are discussed in Chapter 6.

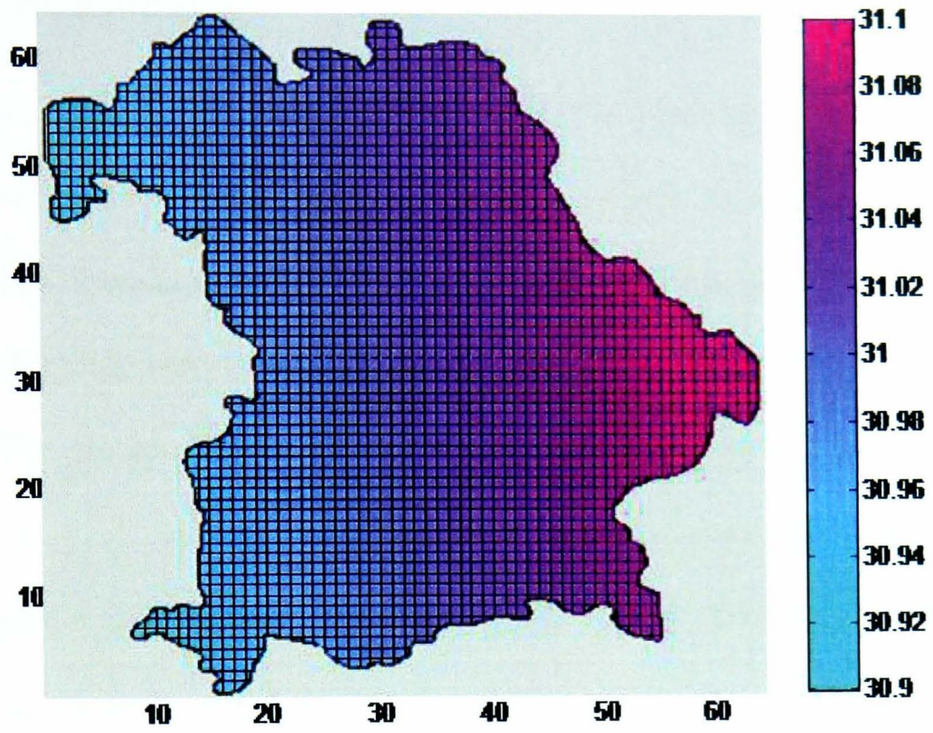


Figure 1.1: Spatial lattice pattern of radioactivity deposition data in Bavaria

1.2.2 Isotropic Spatial Covariances

A process is said to be *isotropic* if the statistical properties of the process are invariant under rotations. For stationary and isotropic processes the covariance function depends only on the scalar distance between two locations s_1 and s_2, namely $\|s_1 - s_2\|$. The covariance is called *anisotropic* when its structure varies with distance and simultaneously as a function of its direction. The proposed model in Chapter 4 can take both *isotropic* and *anisotropic* covariances. But the application to radioactive deposition data in Chapter 5 used the *isotropic* covariance functions.

1.2.3 Stationarity

In general, a process is called *stationary* if all the probability statements about the process within a finite region $\Omega \subset \mathbb{R}$ are invariant under translation of that region. The discussions in this thesis assume that the spatial process has a mean $\mu(s) = E(Y(s))$, associated with it and that the variance of $Y(s)$ exits for all $s \subset D$. The process $Y(s)$ is said to be *Gaussian* if, for any $n \geq 1$ and any set of sites $s_1, \ldots, s_n$, $Y = (Y(s_1), \ldots, Y(s_n))'$ has a multivariate normal distribution. The process is said to be *strictly stationary* if, for any given $n \geq 1$, any set of n sites $s_1, \ldots, s_n$ and any $h \in \mathbb{R}^r$, the distribution of $(Y(s_1), \ldots, Y(s_n))$ is the same as that of $(Y(s_1 + h), \ldots, Y(s_n + h))$. Here $\mathbb{R}^r$ denotes the r-dimensional $(r > 0)$ Euclidean space and D is envisioned as $\mathbb{R}^r$ as well.

A less restrictive condition is given by *weak stationarity* (also called second-order stationarity). Cressie (1993) defines a spatial process to be weakly stationary if $\mu(s) \equiv \mu$ (i.e., the process has a constant mean) and $Cov(Y(s), Y(s+h)) = C(h)$ for all $h \in \mathbb{R}^r$ such that s and $s + h$ both lie within D. Weak stationarity implies that the covariance relationship between the values of the process at any two locations can be summarized by a covariance function $C(h)$, and this function depends only on the separation vector h.

Stationarity requires that the process has constant mean and constant variance. It seems a rather restrictive assumption to require of an environmental process. However, certain types (but not all) of non-stationary processes can be

made stationary through appropriate transformations. More specifically, a spatial process $Z(s)$ is said to be (weakly) stationary if it has constant mean and the spatial covariance depends only on distance between two points. This type of covariance is called Covariogram.

1.2.4 Valid Spatial Covariances

In any spatial modelling, the spatial covariance function should be valid. A spatial covariance function is said to be valid if it is positive definite and rotation invariant, i.e. it is a function which can be used to represent the covariance of isotropic, stationary random fields which we have defined above. According to Schlather (1999), many functions from their respective family are valid in any specified dimension. In the thesis, the proposed HBST model used the spatial covariances from the exponential family and the spherical family. But there exist some other covariance functions which are valid as described by Schlather (1999), some of them are described below.

(a) Bessel family (valid in $\mathbb{R}^d$)

$$\phi_v(h) = 2^v \Gamma(v+1) h^{-v} J_v(h), \qquad v \geq (d-2)/2$$

where J_v is a Bessel function. This function is positive definite in $\mathbb{R}^d$ for $v = (d-2)/2$. The Bessel function with $v = (d-2)/2$ is the isotropic analogue of the cosine in $\mathbb{R}^d$. For arbitrary $v \geq (d-2)/2$, the positive definiteness of the function is demonstrated in Yaglom (1987).

(b) Cauchy family (valid in $\mathbb{R}^\infty$)

$$\phi(h) = (1 + h^2)^{-v}.$$

This function is positive definite as it is the scale mixture of the Gaussian model with mixing density function proportional to

$$f(x) = e^{-x^2} x^{2v-1}.$$

In Cressie (1993), it is called "rational quadratic model" when $v = 1$.

(c) Cosine (valid in $\mathbb{R}^1$)

$$\phi(h) = cos(h).$$

This model is the most elementary covariance model on the real line and it is not a valid model for an isotropic random field in any dimension $d \geq 2$.

(h) Whittle-Matérn class or Basset family (valid in $\mathbb{R}^\infty$)

$$\phi(h) = 2^{1-v}\Gamma(v)^{-1} h^v K_v(h), \qquad v > 0.$$

This class of covariance functions is very popular as it allows the user to specify the degree of differentiability of the underlying random field. Here $K_v(\cdot)$ denotes the modified Bessel function of the third kind of order v.

(g) Symmetric stable family or powered exponential family (valid in $\mathbb{R}^\infty$)

$$\phi(h) = e^{-h^v}, \qquad v \in [0, 2].$$

These models are positive definite.

(d) Exponential family (valid in $\mathbb{R}^\infty$)

$$\phi(h) = e^{-h}.$$

This model is a special case of the Whittle-Matérn family, namely the model with parameter $v = 1/2$ and a special case of the symmetric stable family, namely the model with parameter $v = 1$. The exponential model is also interesting in its own right, as it allows for very simple and fast simulations on the real axis because of its Markov property.

(e) Gaussian family (valid in $\mathbb{R}^\infty$)

$$\phi(h) = e^{-h^2}.$$

This is a special case of the symmetric stable family, namely the model for $v = 2$. In the two-dimensional case, its spectral distribution function is

$$F(x) = 1 - exp(-x^2/4).$$

Thus, a random field with Gaussian model as covariance function can be simulated by the spectral method. Although the Gaussian model is still frequently used, it has two disadvantages.

(i) The covariance matrices involved in simulation studies are often almost singular, and therefore numerically unstable. This kind of behaviour is to be expected as the Gaussian model is on the border of the set of

valid models, as $\phi(h) = exp(-h^v)$ is not a valid model in any dimension for any $v > 2$.

(ii) The random field itself has a strange behaviour: any realisation is completely determined on the entire space by the values within an arbitrarily small, open ball.

(f) Spherical family (valid in $\mathbb{R}^3$)

$$\phi_3(h) = 1 - \frac{3}{2}h + \frac{1}{2}h^3, \qquad 0 \leq h \leq 1.$$

The name "spherical" refers to the fact that $\phi_3(h)$ is the volume of the intersection of two three-dimensional balls with unit diameter and center a distance h apart. For some purposes, it is also convenient that the correlation has a finite range. Because the family depends only on a scale parameter h, it gives no flexibility in shape.

1.2.5 Markov Random Fields

A Gaussian Markov random field is a model for spatially distributed random variables.

Suppose $Z = (Z_1, Z_2, ..., Z_n)'$ have a normal distribution with mean μ and covariance matrix Σ. Define the labelled graph $G = (\gamma, \epsilon)$, where γ ={1,2,.,n} and ϵ be such that there is no edge between node i and j iff Z_i II $Z_j|Z_{-ij}$, that is Z_i and Z_j are conditionally independent, where Z_{-ij} is short for $Z_{-\{i,j\}}$. Then we say that Z is a Gaussian Markov Random Field (GMRF) with respect to G.

Since the mean μ does not have any influence on the pair wise conditional independence properties of Z, we can deduce that this information must be 'hidden' solely in the covariance matrix Σ. It turns out that the inverse covariance matrix, the precision matrix $Q = \Sigma^{-1}$ plays the key role.

1.2.6 Kalman Filter

In 1960, R. E. Kalman published his famous paper (Kalman, 1960) describing a recursive solution to the discrete-data linear filtering problem. The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown.

The Kalman filter is a recursive predictor-corrector type estimator. The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. As such, the equations for the Kalman filter fall into two groups: *time update* equations and *measurement update* equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the *a priori* estimates for the next time step. The measurement update equations are responsible for the feedback - i.e. for incorporating a new measurement into the *a priori* estimate to obtain an

improved *a posteriori* estimate.

The time update equations can also be thought of as *predictor* equations, while the measurement update equations can be thought of as *corrector* equations. Indeed the final estimation algorithm resembles that of a *predictor-corrector* algorithm for solving numerical problems as shown below in Figure 1.2.

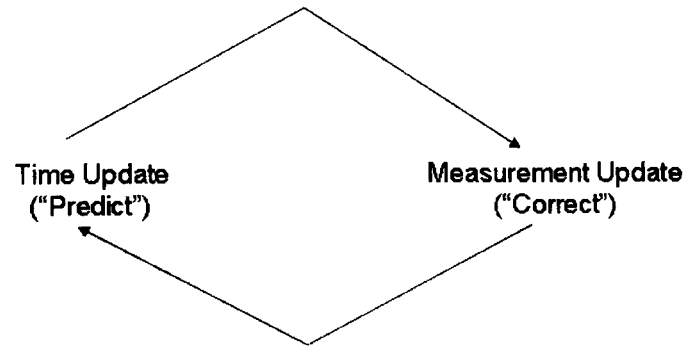


Figure 1.2: The ongoing discrete Kalman filter cycle. The *time update* projects the current state estimate ahead in time. The *measurement update* adjusts the projected estimate by an actual measurement at that time.

After each time and measurement update pair, the process is repeated with the previous *a posteriori* estimates used to project or predict the new *a priori* estimates. This recursive nature is one of the very appealing features of the Kalman filter.

1.2.7 Kriging

In real situation, it is impossible to get all-inclusive values of data at every desired point because of practical constraints. Thus, interpolation is important and fundamental to graphing, analysing and understanding of 2D data. Kriging is a well-known method of spatial interpolation used in geostatistics. It was named

after D. G. Krige from South Africa. It is a method of interpolation which predicts unknown values from data observed at known locations. This method uses variogram (which is a function describing the degree of spatial dependence of a spatial random field, Cressie (1993), pages. 58-59) to express the spatial variation, and it minimizes the error of predicted values which are estimated by spatial distribution of the predicted values. For detail about kriging please see Cressie (1993). Following are the different types of kriging.

- *Ordinary Kriging*: A variety of kriging which assumes that local means are not necessarily closely related to the population mean, and which therefore uses only the samples in the local neighborhood for the estimate. Ordinary kriging is the most common method used in environmental engineering.

- *Block Kriging*: Estimating the value of a block from a set of nearby sample values using kriging.

- *Point Kriging*: Estimating the value of a point from a set of nearby sample values using kriging. The kriged estimate for a point will usually be quite similar to the kriged estimate for a relatively small block centered on the point, but the computed kriging standard deviation will be higher. When a kriged point happens to coincide with a sample location, the kriged estimate will equal the sample value.

- *Universal Kriging*: Kriging similar to that of ordinary kriging, but used when a trend, or slow change in average values, in the samples exists.

1.3 Notation

Throughout the thesis, I adopt the standard convention that random variables are denoted as capital letters, and instantiations of random variables (values) are denoted as lower-case letters. I do not explicitly distinguish between scalars and vectors. The only difference is, vectors are denoted by bold-case letters and the scalars are without bold-case but the same letter as vector.

Throughout the thesis, I denote vectors (which may be random variables) as bold-case letters, to distinguish them from matrices, which are always italics. Also, I will use prime to denote vector or matrix transposition.

A typical notation used throughout this thesis is the following:

$$(\boldsymbol{\Theta}_t|\mathbf{I}_t) \sim N[\mathbf{m}_t, C_t]$$

It means that the conditional distribution of the random vector $\boldsymbol{\Theta}_t$ given the value of $\mathbf{I}_t$ (set of all observations up to time t), has a multivariate normal distribution with mean vector $\mathbf{m}_t$ and the covariance matrix C_t.

Finally, it is worth remarking that the equations are numbered according to the chapters. For instance, equation (4.1) means equation number 1 of the Chapter 4. Further notation will be introduced as necessary in each chapter.

A summary of the most frequently used notation and abbreviations appears in the tables in Appendix D.

1.4 Thesis Structure

The thesis is structured as follows. Chapter 2 shows the related literature review. The basic multivariate DLM theory and some related topics are reviewed in Chapter 3 of this thesis, which is the key chapter about background material. This comprises the general model formulation in Section 3.1, as well as the usual form of analysis. Such analysis include the basic conjugate analysis with *known* observational variance-covariance matrix in Section 3.2, and the normal inverse-Wishart conjugate analysis with *unknown* observational variance-covariance matrix in Section 3.3.

The main results of the thesis are presented through Chapters 4 to 7 as follows. Chapter 4 describes the model formulation of the proposed HBST model. In Section 4.1, a general framework of the HBST model introduced. Section 4.2 introduced the different spatial covariance functions. Section 4.3 describes the different parametric families of the spatial covariance functions when the observational covariance matrix is known, and the corresponding basic conjugate analysis is shown in Section 4.4. Section 4.5 describes the normal inverse-Wishart HBST model with unknown observational covariance matrix.

In Chapter 5 an application of the HBST model to radioactivity deposition which allows for uncertainty handling and data assimilation of ECOSYS is introduced. Section 5.1 describes an overview of the physical model for nuclear emergencies, since the proposed HBST model is a part of that whole chain mod-

ules. In Section 5.2, I describe in detail an application of the HBST model to radioactivity deposition. The deposition module of ECOSYS is briefly introduced in Subsection 5.2.1. Subsection 5.2.2 describes the notations and the model components associated with the HBST model as well as a graphical model that establishes a link between the statistical and the ECOSYS models. In Subsection 5.2.3, the HBST model formulation to radioactivity deposition is described. The Subsections 5.2.4 & 5.2.5 defines the three different formulations of the HBST model to the deposition data. Two of those formulations differ on the functional form of their known observational covariance matrix while the third, a normal-inverted-Wishart model with unknown observational covariance matrix.

In Chapter 6, the results from the data assimilation by the HBST model to radioactivity deposition are described. Section 6.1 describes the radioactivity data required for HBST model. More specifically, in Subsection 6.1.1, the real data of Chernobyl's near-ground radioactive deposition in Bavaria is described and in Subsection 6.1.2, the predicted data from a long range atmospheric dispersal model, called K-model is described. Section 6.2 describes the main results of the application of the three HBST models. This Section starts by describing the values of the input parameters of the HBST in Subsection 6.2.1. The summary of posterior results and its validation against measurements are described in Subsections 6.2.2 and 6.2.3. The overall predictive performances and the radionuclide specific predictive performances of the HBST models are described in Subsections 6.2.4 and 6.2.5. Section 6.3 checks the HBST model sensitivity for

the variation in the input parameters of the K-model and the variation in spatial decay of the spatial covariance function for the fixed functional HBST model. The cross-validation of the HBST model is described in Section 6.4 where the whole data set split into two parts, one is training set and the other is validation set.

Finally, Chapter 7 consists of a general conclusion and possible topics for further research.

Appendix A and Appendix B show the prior-to-posterior updating equations for the HBST models with known and unknown observational covariance matrix respectively. Appendix C shows the Cholesky decomposition & eigenvalue decomposition method for the positive definite matrices. A summary of the notation & abbreviations used in the thesis is presented in Appendix D.

Chapter 2

Related Literature Review

This chapter describes the main results in the literature associated with the formulation of the proposed model in Chapter 4 & 5. The proposed model in Chapter 4 is a hierarchical Bayesian space-time (HBST) model for the spatial data. So the following sections are structured as follows.

Section 2.1 gives a short description of the statistical models for spatial data. Section 2.2 reviewed the hierarchical models. In section 2.3, the spatio-temporal models, both in classical and Bayesian approach that are related to the proposed HBST model are reviewed. The HBST model is an extension of dynamic linear model (DLM) to space, so section 2.4 describes the related literature review on DLM. The HBST model built based on a biological model, called, ECOSYS (Müller and Pröhl, 1993) and it takes the prediction from other atmospheric dispersal model (ADM) as the input. Section 2.5 describes the model ECOSYS and related literature on different ADM.

2.1 Statistical Models for Spatial Data

The statistical modelling of data that occur in space is of great interest in many areas, e.g. in environment, in health, in geology, in astronomy, among many others. The low cost Geographic Information System (GIS) is becoming more and more common with user friendly interfaces. These systems allow the spatial visualization of variables such as individual populations, quality of life indexes and company sales in a region using maps. Besides the visual perception of the spatial distribution of the phenomenon, it is very useful to translate the existing patterns into objective and measurable considerations. The purpose of statistical models for describing spatial variation is to find a representation of the variation of a response variable where covariates are not present in the model. The data, then, consist only of measurements of the variable of interest at specified locations where they are observed.

In many classical approaches to the statistical parametric modelling of spatial variation parameters are assumed to be fixed values and estimated by likelihood methods. For example, Cook and Pocock (1983) fit a parametric (linear) model to the trend, followed by an investigation of residual spatial dependence. Besag (1974) proposed a set of spatial models which are constructed using the Markov random-field approach, where he examined the formulation of conditional probability models for finite system of spatially interacting random variables. In the Markov random-field approach, the conditional probability distribution of any

site depends only on its 'nearest neighbours'. That is, suppose $P(X_{i,j}|$ all other site values) denotes the conditional probability distribution of any random variable X at site (i,j), then $P(X_{i,j}|$ all other site values) depends only upon $X_{i,j}$, $X_{i-1,j}$, $X_{i+1,j}$, $X_{i,j-1}$, $X_{i,j+1}$ for each internal site (i,j). Here, each internal site (i,j) has four neighbours, namely $(i-1,j)$, $(i+1,j)$, $(i,j-1)$ and $(i,j+1)$, which are called 'nearest neighbours' of site (i,j). Later Hill *et al.* (1984) and Clayton and Kaldor (1987) assume conditional independence of the data given the parameters, where the *parameters* are assumed to follow a Markov random field.

2.2 Hierarchical Models

The formal ideas of hierarchical modelling arise from simple probability rules. Some early modelling efforts in the mid 20th century that are clearly hierarchical in nature, are not referred as such, e.g. early non-Bayesian works of Stein (1955) and James and Stein (1960) were influential in the development of hierarchical normal models. For a summary please see the bibliographic note in Gelman *et al.* (1995), pages 154-155. Although the concept is not inherently Bayesian, over time most of the literature has been developed in that context and the best description are most often found in the Bayesian literature (Gelman *et al.*, 1995).

Hierarchical modelling is mainly based on the simple fact from probability that the joint distribution of a collection of random variables can be decomposed into a series of conditional models. That is, if Z, P, Q are random variables.

then we can write their joint distribution in terms of a factorization such as $[Z, P, Q] = [Q|P, Z][P|Z][Z]$, where $[P]$ refers to the distribution of P and $[P|Z]$ refers to the conditional distribution of P given Z. This simple formula is the most important point of hierarchical thinking. For example, for a spatio-temporal process, the joint distribution describes the stochastic behaviour of the process at all spatial locations and all times, including all possible interactions. This can be extremely difficult (if not impossible) to specify for many environmental processes. However, it is often much easier to specify the distribution of the relevant conditional models (e.g., conditioning the process at the present time given the past). In this case, the product of a series of relatively simple conditional models leads to a joint distribution that can be quite involved. For complicated processes in the presence of data, the idea is to approach the problem by breaking it into three primary stages (Berliner, 1996):

Stage 1. Data Model: [data | process; parameters]

Stage 2. Process Model: [process | parameters]

Stage 3. Parameter Model: [parameters]

The first stage is concerned with the observational process or 'data model', which specifies the distribution of the data given the process of interest and parameters that describe the data model. The second stage then describes the process, conditional on other parameters. Finally, the last stage accounts for the uncertainty in the parameters. Ultimately, we are interested in the distribution of the process and parameters updated by the data. We obtain this so-called

'posterior' distribution via Bayes' rule:

[process, parameters | data] $\propto$ [data | process, parameters] [process | parameters][parameters].

This formula serves as the basis for Bayesian hierarchical analysis.

2.3 Spatio-temporal Models

Spatio-temporal models have gained widespread popularity in recent years. One reason for this growth is that diverse areas, such as climatology, ecology, environmental, health, and real estate marketing, are increasingly faced with the task of analysing data that are (i) highly multivariate, with many important predictors and response variables, (ii) geographically referenced, and often presented as maps, and (iii) temporally and spatially correlated, as in longitudinal or other time series structures. Some examples of spatio-temporal data are monitoring of regional environmental contamination, disease mapping and the analysis of satellite data. Increased computational power is another reason for the recent surge in this area. Space-time data sets are often large and therefore require substantial computing resources to fit even simple models.

The spatio-temporal models are often constructed by combining time-series models with variogram-based models from spatial statistics. In time-series context, popular approaches include autoregressive moving average models (Box *et al.*, 1994) for stationary data, and state-space models (West and Harrison, 1997), which allow for non-stationary components such as temporal trends and

seasonality. In the spatial setting, much of the literature revolves around isotropic models (Cressie, 1993). The main objective of these models is prediction or kriging. Various methods also exist for non-stationary spatial processes, where the correlation depends on location as well as distance. Sampson and Guttorp (1992) developed a approach which based on transformations of the co-ordinate system.

2.3.1 Classical Spatio-temporal Models

Classical space-time state-space model approaches have been suggested in Mardia *et al.* (1998) combined kriging and state-space models to model spatial covariance and time dependence. They used iterative numerical optimization to estimate variogram parameters and the Kalman filter to estimate temporal trend fields. Similar classical approaches include Wikle and Cressie (1999) and Huang and Cressie (1996). Early statistical models for space-time data often relied on the assumption of temporal stationarity. For example, the STARMA (Pfeifer and Deutsch (1980a) and Pfeifer and Deutsch (1980b)) and STARMAX (Stoffer, 1986) models were constructed by adding spatial covariance matrices to standard vector autoregressive moving average models. Haslett and Raftery (1989) used similar models but relied on fractional differencing to address long memory features in Irish wind power data. Handcock and Wallis (1994) developed a model with kriging and long memory dependence to study global warming trends in northern USA. Another related approach was taken by Carroll *et al.* (1997), who estimated ground level ozone through Gaussian random fields by assuming separable space

and time correlation functions.

In this field, different types of spatio-temporal models have developed. For instance, Hirst *et al.* (2003) has developed a hierarchical model on long-transported air pollution in Europe. This is basically a spatio-temporal linear regression model for the true deposition at location s and time t. Hirst *et al.* (2003) used two stochastic terms in their model for different spatial covariance of deposition. For example, deposition in mountainous areas might have a high variance and a short correlation range, and in other areas correlation might be greatest in the direction of prevailing winds.

2.3.2 Bayesian Spatio-temporal Models

In last few years, several models have been proposed for non-stationarity and spatial anisotropy in spatiotemporal data. For example, Guttorp *et al.* (1994) used the deformation technique of Sampson and Guttorp (1992) to capture spatial anisotropy in a temporally prewhitened series of ozone readings. Higdon (1999) relied on a Bayesian process convolution approach to model temporal and spatial non-stationarity in north Atlantic Ocean temperatures. Lavine and Lozier (1999) proposed a Bayesian model for ocean temperatures using a three-dimensional Markov random field. Smith and Robinson (1997) proposed a Gibbs sampling approach for estimating rainfall in the presence of storms and Dellaportas *et al.* (1998) developed space-time models for wind speed. Waller *et al.* (1997) used a hierarchical Bayesian model with conditional autoregressive priors for spatial

effects for the mapping of Ohio lung cancer rates. Other approaches involving hierarchical Bayesian models include Wikle *et al.* (1998), who analysed monthly maximum atmospheric temperatures, and Gelfand *et al.* (1998), who estimated a model for real-estate prices.

Another modelling approach for time series is through the state space framework. Sansó and Guenni (1999) assumed a parametric covariance function for the observation errors and fitted a model by using Markov chain Monte Carlo (MCMC) simulation. Tonellato (1997) proposed a vector AR(p) model with spatially correlated time shocks for fixed observation stations and applied it in Tonellato (1998) for wind power prediction. The method of Tonellato (1997) is limited by the fact that the dimension of the state vector grows in the number of observation stations.

One of the main problems in spatial statistical modelling is to determine a spatial covariance matrix which not only represent spatial association in a realistic way but also is a valid one, that is, isotropic, positive definite. Notice that later requirement restricts the choice of valid functions when a functional form is used in the modelling.

Instead of any fixed functional form of the spatial covariance function, we can learn from data on covariance on Bayesian paradigm. Le and Zidek (1992) developed a hierarchical Bayesian alternative to Kriging. Kriging is well-known method of spatial interpolation and it is an important tool in geostatistics. But for environmental aspects, Kriging is less suitable as a tool for interpolating spatial

random fields which are observed successively over time. So, Le and Zidek (1992) presented temporal (and spatial) modelling in their paper which is closely related related to one of our proposed model in chapter 4.

More related to a model formulation used in this thesis, Le and Zidek (1992) proposed a method that used an inverted Wishart prior for the spatial covariance function. They proposed the linear regression model,

$$\mathbf{Y}_t|\mathbf{X}_t, B, \Sigma \sim N_p\big(B\mathbf{X}_t, \Sigma\big)....independently \tag{2.1}$$

Suppose a real-valued spatial random field is observed at g discrete 'monitored' sites and yields a $g \times 1$ data vector, $\mathbf{Y}_t^{(2)} = \{Y_t^{(21)} \dots Y_t^{(2g)}\}$. Let the object of inferential interest is a $u \times 1$ vector, $\mathbf{Y}_t^{(1)} = \{Y_t^{(11)} \dots Y_t^{(1u)}\}$, of 'unmeasured' values at u 'unmonitored' sites. So the spatial field is over $p = u+g$ discrete sites. Now, we partitioned the p dimensional random vector $\mathbf{Y}_t$ as $\mathbf{Y}_t = (\mathbf{Y}_t^{(1)}, \mathbf{Y}_t^{(2)})'$ where $\mathbf{Y}_t^{(1)}$ and $\mathbf{Y}_t^{(2)}$ are the u and g dimensional vectors described above. Let, $\mathbf{X}_t = \{X_{1t} \dots X_{kt}\}$ is a k-dimensional vector of covariates and B is a $(p \times k)$ dimensional matrix of the regression coefficients. $B = (B_1, B_2)'$ and Σ has an inverse Wishart distribution. According to them, the hyperparameters of the second-level priors are specified.

The covariance matrix Σ is partitioned as,

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Now, Σ is reparametrized as $(\Sigma_{22}, \Sigma_{1|2}, V)$, where $\Sigma_{1|2}$ is a $(u \times u)$ matrix and

$\Sigma_{1|2} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$. Here, Σ_{22} is a $(g \times g)$ matrix and $V = \Sigma_{12}\Sigma_{22}^{-1}$

Through the Bartlett decomposition, Σ can be written as $\Sigma = T\Delta T'$, where

$$\Delta = \begin{bmatrix} \Sigma_{1|2} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$$

and

$$T = \begin{bmatrix} I & V \\ 0 & I \end{bmatrix}$$

where 0 is a null-matrix and I is the identity matrix.

Σ can be written as :

$$\Sigma = \begin{bmatrix} \Sigma_{1|2} + V\Sigma_{22}V' & V\Sigma_{22} \\ \Sigma_{22}V' & \Sigma_{22} \end{bmatrix}$$

And, $\Sigma|\Psi, m \sim W_p^{-1}(\Psi, m)$ which is p variate inverse-Wishart distribution with m degrees of freedom.

From the main results obtained by Le and Zidek (1992), the posterior distribution of $\left(\Sigma_{22}, \Sigma_{1|2}, V\right)$ is (here, $p = u + g$ and $\mathbf{I}$ is a set of n independent realizations)

$$\Sigma_{22}|\mathbf{I}, \Psi_{22}, m \sim W_g^{-1}(\hat{\Psi}_{22}, m + n - u) \tag{2.2}$$

$$\Sigma_{1|2}|\mathbf{I}, \Psi_{1|2}, m \sim W_u^{-1}(\Psi_{1|2}, m) \tag{2.3}$$

$$V|\mathbf{I}, \eta, \Sigma_{1|2} \sim N_{ug}(\eta, \Sigma_{1|2} \otimes \Psi_{22}^{-1}) \tag{2.4}$$

Where,

$$\hat{\Psi}_{22} = \Psi_{22} + S + (\hat{B}_2 - B_2^0)'(A^{-1} + F^{-1})^{-1}(\hat{B}_2 - B_2^0) \tag{2.5}$$

and

$$S = \sum_{t=1}^{n} (\mathbf{Z}_t^{(2)} - \hat{B}_2 X_t)(\mathbf{Z}_t^{(2)} - \hat{B}_2 X_t)' \tag{2.6}$$

A spatial model, like any other statistical model, also has assumptions and requirements it must satisfy. Perhaps one of the most common requirements is that the underlying process be stationary.

Another important issue of the spatio-temporal models is associated with their highly dimensional structure. This is the subject of concern because their use require a considerable computational demand.

2.4 The Dynamic Linear Model

The Dynamic Linear Model (DLM), West and Harrison (1997), can be represented as a system of equations specifying how observations of a process are stochastically dependent on the current process state and can be represented by how the process parameters evolve in time. The system of equations captures the inherent process dynamics with the stochastic elements modelled by random shocks or disturbances. The DLM is a general class of linear models flexible enough to represent most real time series processes. The model is stated in terms of discrete, equally spaced intervals of time although it is possible to extend it to

unequal intervals. Cases with missing values can also be dealt with within the DLM formulation, West and Harrison (1997).

The DLM formulation must follow certain principles for Bayesian forecasting and dynamic modelling. Suppose that at a certain time $t-1$ all the relevant information available up to that time is denoted as $\mathbf{I}_{t-1}$, i.e. $\mathbf{I}_{t-1}$ is the set of all observations up to time $t-1$. From the modellers/forecaster point of view interest lies in the forecast value of some vector quantity $\mathbf{Y}_t$ with observed values of this quantity as $\mathbf{y}_t$. It follows that $\mathbf{I}_t = \{\mathbf{Y}_t, \mathbf{I}_{t-1}\}$. At time $t-1$ the adopted parameterisation is that all historical information relevant to predicting future observations $\{\mathbf{Y}_t, \mathbf{Y}_{t+1}, \dots\}$ is contained in a n- dimensional vector denoted θ_{t-1}. This relevant information is represented in terms of a conditional probability distribution $(\theta_{t-1} \mid \mathbf{I}_{t-1})$ such that given $\mathbf{I}_{t-1}$, $(\theta_{t-1} \mid \mathbf{I}_{t-1})$ is sufficient for predicting the future. The parameter vector θ_{t-1} must be meaningful, dynamic and changeable to allow the incorporation of expert information from the decision makers and from other influential factors outside the system. Current information can then be related to the future via some derived predictive distribution $(\mathbf{Y}_{t+k} \mid \mathbf{I}_{t-1})$, $k = 0, 1, 2, \dots$. This derivation is via specification of a sequential parametric relation $(\theta_t \mid \theta_{t-1}, \mathbf{I}_{t-1})$ together with $(\mathbf{Y}_t \mid \theta_{t-1}, \mathbf{I}_{t-1})$. In combination with $(\theta_{t-1} \mid \mathbf{I}_{t-1})$ these distributions enable derivation of a full joint forecast Student t-distribution. The main property enabling effective dynamic modelling is *conditional* independence, which can be stated generally as follows: given the present state θ_t, the present observation $\mathbf{y}_t$ and the future observation

$\mathbf{y}_{t+m}$ for $m \geq 1$ are independent of the past observation $\mathbf{y}_{t-1}$.

For a time point t, $(t = 1, 2, \dots)$, $\mathbf{Y}_t$ is time series with observation $\mathbf{y}_t$. Let θ_t be a r-dimensional vector of parameters at time t. We setup the model by the four quantities $\{F_t, V_t, H_t, W_t\}$, which we assume are known at time t. The model can be expressed as:

$$\mathbf{Y}_t = F_t'\theta_t + \nu_t \qquad \nu_t \sim N(\mathbf{0}, V_t) \tag{2.7}$$

$$\theta_t = H_t\theta_{t-1} + \omega_t \qquad \omega_t \sim N(\mathbf{0}, W_t) \tag{2.8}$$

where F_t is a $n \times r$ matrix, V_t is $r \times r$ variance-covariance matrix, H_t is $n \times n$ matrix and W_t is a $n \times n$ variance matrix. Here V_t and W_t are symmetric and positive (semi)-definite. There are two main interpretations of the process parameters. A process parameter may be expressed as a relationship between dependent variables $\mathbf{Y}_t$ and a matrix of independent variables F_t in which case it extends the classical parametric interpretation of the static model. The second interpretation is if F_t were constant the model is a state space representation of the time series in which the parameters may be thought of as process growth and so on.

Equation (2.7) and (2.8) are not a complete specification, as it does not tell us of the relationship between the quantities ν_t and ω_t, nor the relationship between, say, ν_t and ω_{t-1}. The best way to think of a DLM is as the graphical structure in Figure 2.1 imposed on the model in equations (2.7) and (2.8). This graph has two features. First, the state vector θ follows a markov process, so that the

conditional independence $\theta_k \amalg \theta_j \mid \theta_{k-1}$, for all $1 \leq j < k \leq t$. This implies the marginal independence $\omega_j \amalg \omega_k$ in equation (2.8). Second, θ_k separates $\mathbf{Y}_k$ from everything else, so that the conditional independence $\mathbf{Y}_k \amalg \mathbf{Y}_j, \theta_j \mid \theta_k$. This implies the independence $\nu_k \amalg \omega_k$ and $\nu_k \amalg \nu_j, \omega_j$ in equations (2.7) and (2.8).

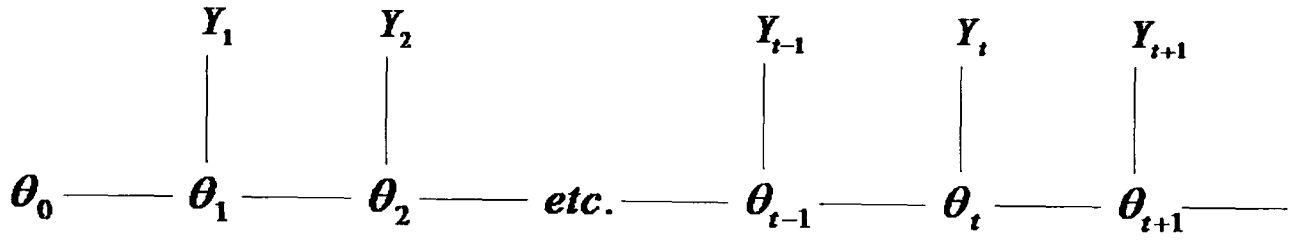


Figure 2.1: The DLM conditional independence structure

A detailed review of the basic multivariate DLM theory and some related topics is given in Chapter 3 as this is the key chapter about background material for the formulation of our proposed model in Chapter 4.

Like Gamerman *et al.* (2006), we adopt a state-space approach based on an extension of the DLM of West and Harrison (1997) to our spatio-temporal process. However, unlike Gamerman *et al.* (2006), the proposed approach in this thesis includes specific environmental and physical models in its formulation. The paper by Gamerman *et al.* (2006) concentrates on the simplest version of the spatio-temporal process and only considers models for continuous space.

Similarly Stroud *et al.* (2001) proposed a state-space model for non-stationary spatio-temporal data. In another development, Shaddick and Wakefield (2002) have proposed a spatio-temporal model for four pollutants measured daily at

eight monitoring sites in London over a 4-year period. They have suggested a hierarchical DLM with three stages. This model allows a temporal-pollutant interaction, and a spatial-pollutant interaction. They have restricted the regression coefficients as scalar, which does not depends on the time component. In our proposed model, the coefficients are depends space as well as time simultaneously. Also, in our proposed model the error terms are not necessarily stationary, it can be non-stationary.

In another development, Lavine (1998) proposed a model based on Markov random field prior on lattice design. He shows that a Markov random field prior on a lattice is identical to the posterior from a particular DLM updated with a particular data set.

Politis and Robertson (2004) proposed a state-space model based on an extension of the DLM of West and Harrison (1997) to predict the dispersal of contamination on a large scale grid in the event of an accidental release of radioactivity. Similarly to Politis and Robertson (2004), our proposed model built on a physical model for atmospheric dispersion and transport. Unlike Politis and Robertson (2004), the state parameters in our proposed model in this thesis are not deterministic.

As mentioned above that the proposed model in Chapter 4 built on a physical model for atmospheric dispersion and transport, so the next section reviewed the physical model, called ECOSYS (Müller and Pröhl, 1993) and the different atmospheric dispersal models.

2.5 Environmental Models for Radioactivity Contamination and Deposition

Environmental models can be broadly divided into short and long range models, depending on the scope and the extent of their application.

The following subsections describe a food chain model, called ECOSYS, developed by Müller and Pröhl (1993), and some atmospheric dispersal models including the K-model by Lauritzen and Mikkelsen (1999) which is used in this thesis as we shall see.

2.5.1 ECOSYS

ECOSYS is a time-dependent radiological simulation model which has been developed to access the radiological consequences of short-term depositions of radionuclides. This model has been developed by Müller and Pröhl (1993). ECOSYS is a dynamic model of radionuclides migration in terrestrial chains that contain soil, plants, agricultural products and humans. This model uses inputs as the integral of activity concentration in air and deposition on ground for each radionuclide and the amount of rainfall. The model includes the 'internal exposure' via inhalation and ingestion and the 'external exposure' from the passing cloud and from radioactivity deposited on the ground. By means of the interception on plants and the deposition on ground the transfer of the radionuclides in different ecosystems and different feed-food chains can be simulated. Finally the course of the activity concentrations in foodstuff is calculated.

The ECOSYS model by Müller and Pröhl (1993) is based on one single radiological region. There are many model parameters that depend on the climatological, agricultural, or other characteristics of the region considered. The parameter values given in Müller and Pröhl (1993) are mainly representative of Southern German conditions. However, the design of ECOSYS supports its adaptation to other agricultural, climatological, and radiological situations enabling its application in different geographical regions. The doses via inhalation and external exposure from the cloud are calculated from the concentration of activity in air. The model starts with estimating deposition to a variety of surfaces (vegetation and soil) for the ingestion pathway and external exposure from the ground. For the ingestion pathway, all relevant transfer processes among soil, plants, animals, and processed products, such as interception, translocation, root uptake, animal feeding, food processing, and culinary preparation are considered. ECOSYS is predominantly a model of migration of radionuclides through the food chain, which is adopted to the landscape conditions, consider physical and chemical properties of soil by means of influence of generalized soil factors on the probabilities of transfer of microelements from soil into plants, and consider season of changes in the probabilities of transfer of microelements from soil into underground crop.

2.5.2 Atmospheric Dispersal Models

Atmospheric dispersal models fall into two broad categories: Lagrangian and Eulerian. Lagrangian models use numerical methods to simulate the trajectories of the particles emitted, to approximate a continuous release. That is, a Lagrangian dispersion model mathematically follows pollution plume particles as the particles move in the atmosphere and they model the motion of the particles as a random walk process. The Lagrangian model then calculates the air pollution dispersion by computing the statistics of the trajectories of a large number of the pollution plume particles. Smith and French (1993) suggested that the simple version of these models cannot provide accurate real-time predictions quickly enough.

Eulerian models, in contrast, use the numerical solution to a more complete transport diffusion equation which describes the contaminated plume. An Eulerian dispersion model is similar to a Lagrangian model in the sense that it also tracks the movement of a large number of pollution plume particles as they move from their initial location. Under certain condition conditions it is sometimes possible to obtain analytic steady-state solutions. The Gaussian plume model (RIMPUFF model described below) is one solution which considers the stationarity, constant wind vector and the homogeneous geographic area. But these solutions are analytic and parameterized form and they are steady-state solutions of a system which are not helpful in the early stage of a release.

Mikkelsen *et al.* (1984) have suggested the puff model which assumes the

continuous emissions from source and the contaminated mass has released as a discrete series of small puffs. These puffs are then transported and dispersed by the wind field. Puffs models are simple to work with and possess some useful properties. First, stationary conditions associated with source emissions can be relaxed because different masses under puffs can reflect the often uneven pattern of accidental releases. Second, the parameters associated with each puff dispersal can be made different, reflecting the characteristics of the wind field at the location of that puff. Päsler-Sauer (1986) have suggested that the puff models work reasonably well in short term releases.

The RIMPUFF (RIsØ Mesoscale PUFF model) is Gaussian puff model developed at Risø (Mikkelsen *et al.* (1984) and Thykier-Nielsen *et al.* (1998)). The RIMPUFF is a Lagrangian mesoscale atmospheric dispersal puff model designed for calculating the concentration and doses resulting from the dispersal of airborne materials. This model can cope well with the instationary and inhomogeneous meteorological situations, which are often of interest in connection with calculations used to estimate the consequences of the short term (accidental) release of airborne materials into atmosphere. The RIMPUFF applies both to homogeneous and inhomogeneous geographic region with moderate topography and instationary meteorological conditions.

ATSTEP is a atmospheric dispersal model developed by RODOS (Päsler-Sauer, 1997). RODOS (Realtime Online DecisiOn Support) is an integrated decision support system, is being developed to aid emergency management in

the event of the accidental release of radioactivity. An important component of RODOS is the prediction of dispersal of contaminated material and evaluation of possible countermeasures. This is a Gaussian puff model for distances up to 50 km. It was developed especially for quick simulation calculations in the case of accidental releases of airborne radioactive contaminants. ATSTEP can calculate real-time diagnoses of the radiological situation during or after a release and dispersion prognoses for 24 hours. The radiological situation is described by the following results calculated with ATSTEP: the concentration in the air near ground (instantaneous and time-integrated), the contamination of ground surface (dry and wet), and the gamma radiation from ground and from the radioactive cloud. These results are presented as time dependent, nuclide specific fields in the whole calculation area in the environment of the source of the release. The following phenomena are considered in the modelling of atmospheric dispersion and the radiological situation in ATSTEP: time dependent meteorology (meteorological tower or SODAR data, forecast data, inhomogeneous wind fields), time dependent nuclide-group specific release rates, thermal energy and rise of the puffs released, dry and wet deposition and corresponding depletion of the cloud, gamma radiation from cloud and from ground, radioactive decay, and potential doses.

ApSimon *et al.* (1985a) developed the Lagrangian puff trajectory model, called MESOS, which was initiated in 1976 to simulate the atmospheric transport and dispersal of radionuclides over distances of several hundred km or more. The

purpose of the MESOS model is to simulate, transport, dispersion and depletion of material released to the atmosphere from point sources, out to distances of 1000 or 2000 km. The MESOS model also uses the real meteorological data extracted from standard observations recorded and reported routinely from meteorological stations across Europe. The purpose is to calculate trajectories of discrete puffs from a given source, and to assess probable conditions in the atmosphere along these trajectories affecting dispersion and depletion in an evolving meteorological situation. In ApSimon *et al.* (1985a), the MESOS model and meteorological data have been used to simulate the dispersion of notional releases of selected nuclide from several European sites and in ApSimon *et al.* (1985b), they discuss the sensitivity of dispersion characteristics and levels of contamination to such factors as site and receptor location, prevailing weather conditions, nuclide deposition and decay rates, and mode of exposure (air contamination, dry and wet deposition). The long-range Lagrangian particle model APOLLO (Atmospheric POLlutant LOng range dispersion model) has been developed by ENEA-DISP (Desiato, 1992), which has been designed to estimate air concentration and ground deposition at distances up to some thousands of kilometers from one point source. Desiato (1992) proposed a sensitivity study of APOLLO with the Chernobyl data.

In our study, we used a long-range atmospheric dispersal model, called K-model (Lauritzen and Mikkelsen, 1999) to provide inputs to our HBST models in the Chernobyl's accident radioactive contamination of Bavaria. The K-model

is discussed in detail in Chapter 6.

Chapter 3

Multivariate Dynamic Linear Model

The previous chapter described the main results in the literature associated to the proposed model in Chapter 4. This chapter provides background material about the multivariate Dynamic Linear Model (DLM) as a precursor to the study of the HBST model in Chapter 4 which is an extension of multivariate DLM to space.

The development of statistical procedures for modelling and analysis of vector time series is a very important theoretical issue with an enormous range of applications in many different areas, e.g. business, engineering, social or natural sciences.

Commonly observed economic processes classified or disaggregated by the geographical region or another factor are generating a multivariate series of data. In such cases the correlation structure among the component series or the joint probability distribution of one subset of series given values of another subset of series can be of extreme importance in a decision making or planning.

The more general assumption of these time processes is non-stationarity. In

natural and Bayesian way it considers the case of dynamic Bayesian models (West and Harrison, 1997) or more specifically multivariate DLMs. Harrison and Stevens (1976) first introduced the multivariate DLM, where the basic theory is the same as univariate DLM, provided the observational error variance-covariance matrix is known. But in practice, the observational error variance-covariance matrix is not known.

In the multivariate DLM analysis with known observational error variance-covariance matrix, the recursive filter, called the Kalman filter, is used for prior to posterior updating. If the observational error variance-covariance matrix is unknown then for specific models a conjugate analysis obtains, but for the general multivariate DLM no such analysis is possible (Barbosa and Harrison, 1992). Such basic initial ideas as well as the introduction of the notation will be considered in the following sections of this chapter.

The formulation of the multivariate DLM as well as the simplest form of analysis for such a model is presented in section 3.1. The problem of specification of the error variances as well as updating algorithm is discussed in section 3.2. In section 3.3, an approximate Normal-inverse Wishart conjugate analysis is discussed with the updating algorithm.

3.1 The General Multivariate DLM

This model has developed by Harrison and Stevens (1976) and is stated in terms of discrete intervals of time indexed by t.

For a time point t, $(t = 1, 2, \dots)$, $\mathbf{Y}_t$ is a time series with observations $\mathbf{y}_t$. Let $\mathbf{\Theta}_t$ be a n-dimensional vector of parameters at time t. We set up the model by the four quantities $\{F_t, V_t, H_t, W_t\}$, which are assumed to be known at time t. The model can be expressed as follows.

The *observation equation* is

$$\mathbf{Y}_t = F_t'\mathbf{\Theta}_t + \nu_t \qquad \nu_t \sim N(\mathbf{0}, V_t) \tag{3.1}$$

The *time evolution equation* is

$$\mathbf{\Theta}_t = H_t\mathbf{\Theta}_{t-1} + \omega_t \qquad \omega_t \sim N(\mathbf{0}, W_t) \tag{3.2}$$

The *prior information* is

$$\mathbf{\Theta}_{t-1}|\mathbf{I}_{t-1} \sim N(\mathbf{m}_{t-1}, C_{t-1}) \tag{3.3}$$

where $\mathbf{I}_{t-1}$ is the set of all observations up to time $t-1$ and $N(\mathbf{m}_{t-1}, C_{t-1})$ denotes the multivariate normal distribution with mean $\mathbf{m}_{t-1}$ and variance C_{t-1}.

F_t is a $n \times r$ matrix of (*known*) constants and / or independent regressors and H_t is $n \times n$ a *known* evolution matrix. The errors $\{\nu_t\}$ and $\{\omega_t\}$ are the sequences of independent zero mean normal random vectors. Also assumed that they are independent of each other. V_t is a $r \times r$ (*known*) observational error variance-covariance matrix and W_t is a $n \times n$ (*known*) evolution error variance-covariance matrix. That is, V_t and W_t are symmetric and positive (semi)-definite.

There are two main interpretations of the process parameters. The first interpretation is, a process parameter may be expressed as a relationship between

dependent variables $\mathbf{Y}_t$ and a matrix of independent variables F_t in which case it extends the classical parametric interpretation of the static model. The second interpretation is if F_t were constant the model is a state space representation of the time series in which the parameters may be thought of as process growth and so on.

The normality assumptions for the error terms ν_t and ω_t are not restrictive since the same estimated quantities are obtained with or without normality assumptions as we shall see.

3.2 Basic Conjugate Analysis : V_t is known

This section describes the analysis of multivariate DLM supposing that V_t is known. Suppose that at a certain time $t-1$ all the relevant information available up to that time is denoted as $\mathbf{I}_{t-1}$, i.e. $\mathbf{I}_{t-1}$ is the set of all observations up to time $t-1$. The updating equations for the process parameter $\mathbf{\Theta}_t$ are using normal theory.

Theorem 3.1. *For the multivariate DLM, one-step forecast and posterior distribution are given, for each t, as follows.*

(a) *Prior at time t* :

$$(\mathbf{\Theta}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{a}_t, R_t) \tag{3.4}$$

where $\mathbf{a}_t = H_t\mathbf{m}_{t-1}$ and $R_t = H_t C_{t-1} H_t' + W_t$

(b) *One-step forecast* :

$$(\mathbf{Y}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{f}_t, Q_t) \tag{3.5}$$

where $\mathbf{f}_t = F'\mathbf{a}_t$ and $Q_t = F_t' R_t F_t + V_t$

(c) *Posterior at time t* :

$$(\mathbf{\Theta}_t|\mathbf{I}_t) \sim N(\mathbf{m}_t, C_t) \tag{3.6}$$

where $\mathbf{m}_t = \mathbf{a}_t + A_t\mathbf{e}_t$ and $C_t = R_t - A_t Q_t A_t'$,

with $A_t = R_t F_t Q_t^{-1}$, and $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$

Proof. Please see the proof of Theorem 16.1 in West and Harrison (1997). □

Note on *Filtering* : The use of current data to revise inferences about previous values of the process parameter is called *filtering*. This is an important tool for retrospective time series analysis where the information recently obtained if filtered back to previous time points. The distribution of $(\mathbf{\Theta}_{t-k}|\mathbf{I}_t)$ for $k \geq 1$ and any fixed t, is called the k-step filtered distribution for the state vector at that time.

3.3 Normal-inverse Wishart approximate Conjugate Analysis : V_t is unknown

In most practical situations, the observational error variance-covariance matrix V_t is not known. If V_t is *unknown* then for specific models a conjugate analysis obtains, but for the general multivariate DLM no such analysis is possible. A precise formulation of DLM's in closed form (exact conjugate prior analysis) is

possible only for *univariate DLM* (West and Harrison, 1997) and also for special very restrictive *common components DLM* (see Quintana (1985), Quintana and West (1987) and West and Harrison (1997)), but they are not generally applicable to general multivariate DLM.

Barbosa and Harrison (1992) proposed an approximate conjugate analysis based on a simple analytic approximation. Their proposed approach is an approximate sequential procedure for multivariate DLM, which is also true for the univariate DLM and the common component DLM.

The basic approximation is based on the factorisation of V_t, so that $V_t = S_t^2$. The factorisation of V_t through Cholesky decomposition can be computationally useful but lacks of uniqueness properties such as invariance to permutations of the corresponding elements of $\mathbf{Y}_t$. So, their proposed method considered all variance-covariance matrix factorisation based on eigenvalues decomposition. (See Appendix C for details and references of Cholesky and eigenvalues decompositions). The updating equations are described below.

Theorem 3.2. *For the multivariate DLM with unknown observational variance-covariance matrix V, the posterior distribution and the one-step forecast are given, for each t, as follows.*

Let V be the unknown constant variance-covariance matrix and V is factorized as $V = S^2$. Let the initial point estimate of $V_0 = S_0^2$ be a reference point. Define the parameter vector $\mu_t = \mathrm{E}(\mathbf{Y}_t|\mathbf{\Theta}_t) = F_t'\mathbf{\Theta}_t$

Step-1 *Time updating* :

$$(\boldsymbol{\Theta}_t|\, V, \mathbf{I}_{t-1}) \sim N(\mathbf{a}_t, R_t) \tag{3.7}$$

where $\mathbf{a}_t = H_t \mathbf{m}_{t-1}$ and $R_t = H_t C_{t-1} H_t' + W_t$

There is no time updating for V, it is supposed to be constant, and the time evolution for $\boldsymbol{\Theta}_t$ is the same as in the standard multivariate DLM model described in Theorem 3.1.

Step-2 *Reparametrization and scaling* :

$$(\mu_t|\, V, \mathbf{I}_{t-1}) \sim N(\mathbf{f}_t, R_t^*) \tag{3.8}$$

where $\mathbf{f}_t = F_t' \mathbf{a}_t$ and $R_t^* = S_c \Sigma_t S_c'$ with $\Sigma_t = F_t' R_t F_t$

where initially the scaling matrix S_c is set up as an identity matrix and updated in step-4.

Step-3 One-step forecast :

$$(\mathbf{Y}_t|\mathbf{I}_{t-1}) \sim T[\mathbf{f}_t, S_{c-1} Q_t S_{c-1}, d_{t-1}] \tag{3.9}$$

where $\mathbf{f}_t = F' \mathbf{a}_t$ and $Q_t = \Sigma_t + I$

and $S_{c-1} = S_{t-1} S_0$ with $S_{t-1} = (\Psi_{t-1}/d_{t-1})^{1/2}$

where S_0 is a reference matrix set up initially such that the initial scale factor S_{c-1} is the identity matrix.

and d_{t-1} is degrees of freedom at time $t-1$.

Step-4 *Observation updating* :

(a) Posterior distribution for μ_t :

$$(\mu_t | V, \mathbf{I}_t) \sim N(\mathbf{m}_t^*, C_t^*) \tag{3.10}$$

where $\mathbf{m}_t^* = \mathbf{f}_t + A_t^* \mathbf{e}_t$ and $C_t^* = R_t^* - A_t^* Q_t^* A_t^{*\prime}$

with $A_t^* = R_t^* Q_t^{*-1}$ and $Q_t^* = R_t^* + I$, and $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$

(b) Posterior distribution for V :

$$(V | \mathbf{I}_t) \sim W^{-1}(\Psi_t, d_t) \tag{3.11}$$

where $\Psi_t = \Psi_{t-1} + J_t J_t'$ and $d_t = d_{t-1} + 1$

with $J_t = S_{t-1}[(Q_t^*)^{1/2}]^{-1} \mathbf{e}_t$ and $S_t = (\Psi_t / d_t)^{1/2}$

and $S_c = S_t S_0$

where S_0 is a reference matrix set up initially such that the initial scale factor S_c is the identity matrix

Step-5 *Inverse reparametrization and scaling* :

$$(\mathbf{\Theta}_t | V, \mathbf{I}_t) \sim N(\mathbf{m}_t, C_t) \tag{3.12}$$

where $\mathbf{m}_t = [\mathbf{a}_t + A_t(\mathbf{m}_t^* - \mathbf{f}_t)] = [\mathbf{a}_t + A_t A_t^* \mathbf{e}_t]$

and $C_t = [R_t + A_t(C_t^{**} - \Sigma_t) A_t'] = [R_t - A_t S_c^{-1} A_t^* Q_t^* A_t^{*\prime} S_c^{-1\prime} A_t']$

with $A_t = R_t F_t \Sigma_t^{-1}$ and $C_t^{**} = S_c^{-1} C_t^* S_c^{-1\prime}$

Proof. Please see Barbosa (1989). □

In the next chapter, we will see the general formulation of the proposed HBST model which is an extension of multivariate DLM to space.

Chapter 4

A Hierarchical Bayesian Space-Time (HBST) Model

This chapter describes the general form of the proposed Hierarchical Bayesian Space-Time (HBST) model. Section 4.1 introduces the basic notation involved and the general framework of the proposed model. Section 4.2 introduces the different spatial covariance functions, which gives an idea of the different formulations of the HBST model when in the first case, the mean and observational covariance matrix are *known* and *fixed* and in the second case both are *unknown.* Section 4.3 describes the different parametric families of the spatial covariance functions when the observational covariance matrix is known. Section 4.4 defines a HBST model with fixed functional observational covariance matrix, and the corresponding basic conjugate analysis describes as we shall see. Section 4.5 defines another variation of a HBST model with a normal-inverse Wishart conjugate prior attributed to the *unknown* mean and the *unknown* observational covariance matrix associated with the model.

4.1 The General Framework of the HBST Model

In this section, the multivariate normal HBST model is defined for a vector of observations $\mathbf{Y}_t = (Y_{1t}, Y_{2t}, \ldots, Y_{(m+n)t})'$, where Y_{it} is the observation at the i^{th} location at time t. Suppose $m+n$ be the total number of location points in a plane ($\mathbb{R}^2$).

A multivariate normal HBST model is an extension of a multivariate normal DLM (West and Harrison, 1997) to space. Suppose, for $t = 1, 2, \ldots, \tau$, where τ is any fixed arbitrary time, the observational vector $\mathbf{Y}_t$ has elements associates with spatial locations, i.e. Y_{it} is the underlying observation at a location i in a plane ($\mathbb{R}^2$). A HBST model for $\mathbf{Y}_t$ is defined by the set

$$\{F, G, H, V, \Sigma, W\}_t = \{F_t, G_t, H_t, V_t, \Sigma_t, W_t\}$$

for each time t, where

(a) F_t is a known $(m+n) \times (m+n)$ dynamic regression matrix;

(b) G_t is a known $(m \times n)$ spatial interpolation matrix;

(c) H_t is a known $(n \times n)$ state evolution matrix;

(d) V_t is a known $(m+n) \times (m+n)$ observational error covariance matrix;

(e) Σ_t is a known $(m \times m)$ spatial interpolation error covariance matrix.

(f) W_t is a known $(n \times n)$ evolution error covariance matrix.

Note that, the spatial interpolation matrix G_t depends on time t only when the additional observation comes in the next time point.

Now, for a regular grid with n points over the geographic region of interest and for m off-grid points representing measuring station sites, we can define $\mathbf{Y}_t = (\mathbf{Y}_{Dt}, \mathbf{Y}_{Mt})'$ and $\mathbf{\Theta}_t = (\mathbf{\Theta}_{Dt}, \mathbf{\Theta}_{Mt})'$ as the observational and the associated state vectors respectively, where the subscripts D and M indicate the spatial location of data it represents, that is, D represents grid points and M measuring sites.

The HBST model is a 3-stage hierarchical model where the first stage links the data vector (observations) with the unobserved state vector, the second stage spatially interpolates the values at measuring sites from the grid points through specified spatial interpolation matrix, and the third stage links the state vector at time t with the state vector at time $t-1$ through a specified time-evolution matrix.

Let, $\mathbf{Y}_{Dt} = (Y_{D1t}, \ldots, Y_{Dnt})'$ and $\mathbf{Y}_{Mt} = (Y_{M1t}, \ldots, Y_{Mmt})'$ be the vectors of observations at grid points and measuring station sites (with dimensions n and m) respectively. Also, let $\mathbf{\Theta}_{Dt} = (\theta_{D1t}, \ldots, \theta_{Dnt})'$ and $\mathbf{\Theta}_{Mt} = (\theta_{M1t}, \ldots, \theta_{Mmt})'$ be the state vectors associated with the observational vectors $\mathbf{Y}_{Dt}$ and $\mathbf{Y}_{Mt}$ respectively, where $\mathbf{\Theta}_{Dt}$ and $\mathbf{\Theta}_{Mt}$ are vectors with elements θ_{Dit} ($i = 1, \ldots, n$ for grid points) and θ_{Mit} ($i = 1, \ldots, m$ for measuring station sites) interpreted as the real *unknown* state parameters at location i on time t. We can now define the HBST model by the following equations.

- The *observation equation* is

$$\mathbf{Y}_t = F_t' \mathbf{\Theta}_t + \nu_t, \qquad \nu_t \sim N[\mathbf{0}, V_t] \tag{4.1}$$

where the observational error terms associated with grid points, and monitoring stations is $\nu_t = (\nu_{1t}, \ldots \nu_{(m+n)t})'$, where ν_{it} is the error associated with the i-th observation of $\mathbf{Y}_t$. The observational error vector is assumed to have a normal distribution with mean vector $\mathbf{0}$ and covariance matrix V_t.

- The *interpolation equation* is:

$$\mathbf{\Theta}_{Mt} = G_t \mathbf{\Theta}_{Dt} + \epsilon_t, \qquad \epsilon_t \sim N[\mathbf{0}, \Sigma_t] \tag{4.2}$$

where G_t is a $m \times n$ spatial interpolation matrix with fixed elements g_{ij} (defined below) and ϵ_t is the vector of spatial interpolation errors assumed normally distributed with $\mathbf{0}$ mean vector and covariance matrix Σ_t.

- The *time evolution equation* is

$$\mathbf{\Theta}_{Dt} = H_t \mathbf{\Theta}_{Dt-1} + \omega_t, \qquad \omega_t \sim N[\mathbf{0}, W_t] \tag{4.3}$$

where H_t is a $n \times n$ time evolution matrix (described below), and ω_t is the vector of time evolution errors, assumed normally distributed with $\mathbf{0}$ mean vector and covariance matrix W_t.

- The *initial information* is

$$(\mathbf{\Theta}_0 \mid \mathbf{I}_0) \sim N[\mathbf{a}_0, R_0] \tag{4.4}$$

for some initial prior moments $\mathbf{a}_0$ and R_0 which values are chosen by the analyst. Here $\mathbf{I}_0$ is a vector of observations at time $t = 0$, where $\mathbf{I}_t = (\mathbf{Y}_t, \mathbf{I}_{t-1})$ represents the information available up to time t.

The observational, space interpolation and time evolution error vectors ν_t, ϵ_t and ω_t are assumed independent and mutually independent respectively, and are independent of $(\mathbf{\Theta}_0 \mid \mathbf{I}_0)$. The observation equation (4.1) relates the $(m+n) \times 1$ data vector $\mathbf{Y}_t$ to the unobserved $(m+n) \times 1$ state vector $\mathbf{\Theta}_t$. The evolution equation (4.3) links the state $\mathbf{\Theta}_{Dt}$ at time t with the state at $t-1$, $\mathbf{\Theta}_{Dt-1}$ through the specified time evolution matrix H_t.

The vector of state parameter at grid points $\mathbf{\Theta}_t$ is considered to be *unknown* and a normal prior probability distribution is attributed to it and the observational covariance matrix V_t in equation (4.1) can be either *known* or *unknown*. In one case, V_t is assumed to be known and has its values fixed, while in the other case, V_t is assumed *unknown* and has a prior inverse-Wishart distribution attributed to it. The first case originates the conjugate normal model while the second case originates the conjugate normal-inverse-Wishart model.

Now, for $t = 1, 2 \ldots \tau$, V_{Dt} and V_{Mt} are the covariance matrices of the observational errors associated with the HBST model for grid points and measuring station sites respectively. Note that, V_{Dt} is not diagonal matrix, but, due to the Markovian structure of our model, V_{Mt} is a diagonal matrix.

Due to the Markovian structure of the interpolation matrix, the interpolation error covariance matrix Σ_t is set as a diagonal matrix with the values of the

diagonal elements as $\sigma_i^2 = \sigma_I^2 d_i^*$ where $d_i^* = \frac{d_i}{|d_{max}^*|}$ with d_i being the distance of the measuring station site i to its nearest grid point and d_{max}^* being the maximum distance over all distances from off-grid points to their nearest grid points.

The elements of the spatial interpolation matrix G_t are chosen to have an exponential form, such that

$$g_{ij} = \begin{cases} \alpha_{ij} exp(-\beta_{ij} d_{ij}), & j \in Ne(i) \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

where $Ne(i)$ denotes the set of nearest adjacent grid points to i. E.g. Figure 4.1 shows the 4×4 regular grid, i represents the measuring station (denoted by 'black circle') and j's are the nearest adjacent grid points to the measuring station i.

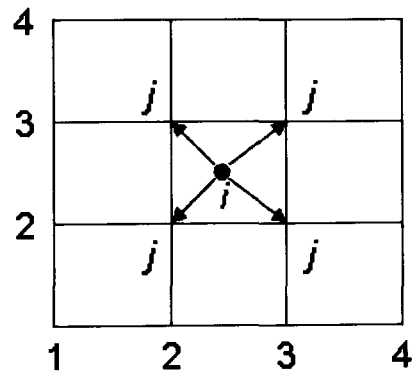


Figure 4.1: The graphical representation of the nearest adjacent grid points to a measuring station i.

also in the equation (4.5), α_{ij} is the proportionality constant such that $\sum_j g_{ij} = 1$, $\beta_{ij} \geq 0$ is the rate of exponential decay with the distance d_{ij}. Note that this choice is a fairly reasonable one in applications (like in chapters 5 and 6) where not only the involved distances are considerably large (as we shall see) but also commonly non-uniform meteorological conditions are largely influential. The

adopted exponential form allows the choice of differing spatial decay rates to reflect particular spatial influences at distinct locations in the underlying region for every component of the HBST model. Those could be chosen for example by experts to reflect prevailing terrain structures at those locations.

The deterministic component of the system variables related to their evolution in time is described by the evolution matrix H_t in equation (4.3) which associates $\mathbf{\Theta}_{Dt}$ with $\mathbf{\Theta}_{Dt-1}$. This association is indirectly carried out through the interpolation equation (4.2). The time evolution matrix H_t is set as a diagonal matrix which diagonal elements has chosen as an exponential form, such that $h_t(i) = \exp(-\lambda_i t)$, where λ_i is the decay parameter.

4.2 The Spatial Covariances of the HBST Model

The spatial covariance matrices associated with the HBST model are V_t, Σ_t and W_t. It is important to set all the spatial covariance matrices for HBST model. The spatial covariance matrices of the HBST model have to be *valid*, i.e. positive definite and rotation invariant. In general, any valid spatial covariance matrices which are invertible can be used by the HBST model for data assimilation. The data assimilation is performed with the use of algorithms in Appendix A which are an extension of the Kalman filter to space-time processes.

Schlather (1999) reviewed many families of the covariance functions which are positive definite and rotation invariant, i.e. valid (as discussed in Chapter 1). There are two cases concerning the values of the observational covariance matrix

V_t. In one case V_t can be known and the values are fixed, while in the other it can be unknown and a prior distribution is attributed to it. These two cases are discussed below.

4.3 Parametric Families of Spatial Covariance Functions of the HBST Model

According to Schlather (1999), a large number of functional covariance forms are available for the elements of *valid* (i.e. invertible) spatial covariance matrices (as discussed in Chapter 1).

Suppose $\rho(d_{ij})$ be the spatial correlation function of the HBST model, where d_{ij} is the distance between two sites i and j. The class of correlation $\rho(d_{ij})$ must be positive definite function. This condition gives the non-negative variance for the covariance function. Checking whether a given function $\rho(d_{ij})$ is positive definite is not straightforward in practice. An equivalent condition is that the spectrum, defined as the Fourier transformation of the covariance function, is non-negative. For further discussion, see Stein (1999) and Schlather (1999).

Apart from the necessary constraint of positive-definiteness, we would usually require the model for the correlation function $\rho(d_{ij})$ to incorporate the following features:

(a) $\rho(d_{ij})$ is monotone non-increasing in d_{ij} (the correlation between two measurements decreases with increasing distance between the two corresponding sampling locations).

(b) $\rho(d_{ij}) \to 0$ as $d_{ij} \to \infty$ (correlation decays to zero at large separation distances).

(c) At least one parameter in the model controls the rate at which $\rho(d_{ij})$ decays to zero (since the separation distance at which the correlation becomes negligible will usually not be known in advance).

Additionally we may wish to include in the model some flexibility in the overall shape of the correlation function. Hence, a parametric model for the correlation function can be expected to have one or two parameters. We now examine briefly some candidate families which meet these requirements.

4.3.1 The Spherical Family

This one-parameter family of correlation functions is defined by

$$\rho(d_{ij};\phi) = \begin{cases} 1 - \frac{3}{2}(d_{ij}/\phi) + \frac{1}{2}(d_{ij}/\phi)^3, & 0 \le d_{ij} \le \phi \\ 0, & d_{ij} > \phi \end{cases} \tag{4.6}$$

The name refers to the fact that $\rho(d_{ij};\phi)$ has a geometric interpretation as the volume of intersection of two three-dimensional spheres whose centres are a distance d_{ij} apart. For some purposes, it is also convenient that the correlation has a finite range. Because the family depends only on a scale parameter ϕ, it gives no flexibility in shape.

4.3.2 The Powered Exponential Family

This two-parameter family is defined by

$$\rho(d_{ij}) = exp\{-(d_{ij}/\phi)^{\kappa}\}, \qquad \phi > 0 \quad \& \quad 0 < \kappa \le 2. \tag{4.7}$$

The exponential correlation function corresponds to the case where $\kappa = 1$. The case $\kappa = 2$ is sometimes called the Gaussian correlation function.

This family often gives a qualitatively reasonable fit to the empirical correlation structure of spatial data.

4.3.3 The Matérn Family

This family is named after Bertil Matérn, who introduced it in his 1960 Swedish doctoral thesis, subsequently reprinted as Matérn (1986). It is defined by

$$\rho(d_{ij}; \phi, \kappa) = \{2^{k-1}\Gamma(\kappa)\}^{-1}(d_{ij}/\phi)^{\kappa}K_{\kappa}(d_{ij}/\phi) \tag{4.8}$$

where (ϕ, κ) are parameters and $K_{\kappa}(\cdot)$ denotes the modified Bessel function of the third kind of order κ.

The family is valid for $\phi > 0$ and $\kappa > 0$. The case $\kappa = 0.5$ is the same as the exponential correlation function, $\rho(d_{ij}) = exp(-d_{ij}/\phi)$. The Gaussian correlation function is the limiting case as $\kappa \to \infty$.

Also, according to Berger *et al.* (2001), the standard families of valid covariance functions are Power Exponential (De Oliveira *et al.*, 1997), Spherical (Wackernagel, 1995), Rational Quadratic (Yaglom, 1987) and Matérn (Matérn (1986); Handcock and Stein (1993)) families.

Although the existence of a large number of functional covariance forms available for the elements of valid (i.e. invertible) spatial covariance matrices, I have chosen two particular forms to define two variations of the HBST model. One has an exponential form proposed by Cressie (1993) and the other has a spherical

form proposed by Schlather (1999) as described in Chapter 5.

In order to operate the models, we need at the initial time $t = 0$ (before any observation is taken) not only to specify the parameters associated with the HBST model formulation, but also the observational, the interpolation and the time evolution covariance matrices (i.e. V_0, Σ_0 and W_0). After those values are stipulated, data assimilation is performed with the use of the algorithms in Appendix A which are an extension of the Kalman filter to space-time processes.

4.4 Basic Conjugate Analysis for *Known* Observational Covariance Matrix

This subsection describes the analysis of HBST model with *known* observational covariance matrix V_t. Suppose that at a certain time $t-1$ all the relevant information available up to that time is denoted as $\mathbf{I}_{t-1}$, i.e. $\mathbf{I}_{t-1}$ is the set of all observations up to time $t-1$. The updating equations for the process parameter $\mathbf{\Theta}_t = (\mathbf{\Theta}_{Dt}, \mathbf{\Theta}_{Mt})'$are using normal theory.

Theorem 4.1. *For the HBST model, one-step forecast and posterior distribution are given, for each t, as follows.*

Let for some mean $\mathbf{m}_{t-1}$ and variance-covariance matrix C_{t-1} :

$$(\mathbf{\Theta}_{Dt-1}|\mathbf{I}_{Dt-1}) \sim N(\mathbf{m}_{t-1}, C_{t-1}) \tag{4.9}$$

where $\mathbf{I}_{Dt-1} = \{\mathbf{Y}_{Dt-1}, \mathbf{I}_{Dt-2}\}$

Let for some mean $\mathbf{a}^*_{t-1}$ and variance-covariance matrix R^*_{t-1} :

$$(\mathbf{\Theta}_{t-1}|\mathbf{I}_{t-1}) \sim N(\mathbf{a}^*_{t-1}, R^*_{t-1}) \tag{4.10}$$

(a) *Prior at time t* :

$$(\mathbf{\Theta}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{a}_t, R_t) \tag{4.11}$$

where $\mathbf{a}_t = \begin{pmatrix} H_t\mathbf{m}_{t-1} \\ \\ G_tH_t\mathbf{m}_{t-1} \end{pmatrix}$

and $R_t = \begin{pmatrix} H_tC_{t-1}H_t' + W_t & G_tH_tC_{t-1}H_t' + G_tW_t \\ \\ H_tC_{t-1}H_t'G_t' + W_tG_t' & G_tH_tC_{t-1}H_t'G_t' + G_tW_tG_t' + \Sigma_t \end{pmatrix}$.

(b) *One-step forecast* :

$$(\mathbf{Y}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{f}_t, Q_t) \tag{4.12}$$

where $\mathbf{f}_t = F'\mathbf{a}_t$ and $Q_t = F_t'R_tF_t + V_t$, with $\mathbf{a}_t$ and R_t as described in equation (4.11).

(c) *Posterior at time t* :

$$(\mathbf{\Theta}_t|\mathbf{I}_t) \sim N(\mathbf{a}^*_t, R^*_t) \tag{4.13}$$

where $\mathbf{a}^*_t = \mathbf{a}_t + A_t\mathbf{e}_t$ and $R^*_t = R_t - A_tQ_tA_t'$,

with $A_t = R_tF_tQ_t^{-1}$, and $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$

where $\mathbf{Y}_t = (\mathbf{Y}_{Dt}, \mathbf{Y}_{Mt})'$ and $\mathbf{f}_t = F'\mathbf{a}_t$ as described in equation (4.12).

Proof. The proof is by induction using the multivariate normal distribution theory. By using the terms in theorem statement, we can write the joint distribution of $(\mathbf{\Theta}_t, \mathbf{Y}_t)$ given by the available information at time $t-1$, $\mathbf{I}_{t-1}$ is as follows,

$$\begin{pmatrix} \boldsymbol{\Theta}_t \\ \mathbf{Y}_t \end{pmatrix} \mid \mathbf{I}_{t-1} \sim N\left[\begin{pmatrix} \mathbf{a}_t \\ \mathbf{f}_t \end{pmatrix}, \begin{pmatrix} R_t & A_t Q_t \\ Q_t A_t' & Q_t \end{pmatrix}\right] \tag{4.14}$$

The following results are derived from the basic facts concerning means and variance matrices of linear functions of normal random vectors.

Suppose the statement in equation (4.9) is true, then

(a) From the evolution equation in (4.3) and equation in (4.9) and (4.10) lead to

$$(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{a}_t, R_t)$$

where

$$\begin{aligned} \mathbf{a}_t = E\left(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1}\right) &= E\begin{pmatrix} \boldsymbol{\Theta}_{Dt} \\ \boldsymbol{\Theta}_{Mt} \end{pmatrix} \mid \mathbf{I}_{t-1}\Big) = E\begin{pmatrix} H_t\boldsymbol{\Theta}_{Dt-1} \\ G_t\boldsymbol{\Theta}_{Dt} \end{pmatrix} \mid \mathbf{I}_{t-1}\Big) \\ &= E\begin{pmatrix} H_t\boldsymbol{\Theta}_{Dt-1} \\ G_t H_t\boldsymbol{\Theta}_{Dt-1} \end{pmatrix} \mid \mathbf{I}_{t-1}\Big) \\ &= \begin{pmatrix} H_t\mathbf{m}_{t-1} \\ G_t H_t\mathbf{m}_{t-1} \end{pmatrix} \end{aligned} \tag{4.15}$$

and

$$R_t = \begin{pmatrix} Var(\boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1}) & Cov(\boldsymbol{\Theta}_{Dt}, \boldsymbol{\Theta}_{Mt}|\mathbf{I}_{t-1}) \\ Cov(\boldsymbol{\Theta}_{Mt}, \boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1}) & Var(\boldsymbol{\Theta}_{Mt}|\mathbf{I}_{t-1}) \end{pmatrix} \tag{4.16}$$

where

$$\begin{aligned} Var(\boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1}) &= H_t Var(\boldsymbol{\Theta}_{Dt-1}|\mathbf{I}_{t-1})H_t' + Var(\omega_t) \\ &= H_t C_{t-1} H_t' + W_t \end{aligned} \tag{4.17}$$

$$\begin{aligned} Var(\boldsymbol{\Theta}_{Mt}|\mathbf{I}_{t-1}) &= G_t Var(\boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1})\, G_t' + Var(\epsilon_t) \\ &= G_t(H_t C_{t-1} H_t' + W_t)\, G_t' + \Sigma_t \\ &= G_t H_t C_{t-1} H_t' G_t' + G_t W_t G_t' + \Sigma_t \end{aligned} \tag{4.18}$$

$$\begin{aligned} Cov(\boldsymbol{\Theta}_{Dt}, \boldsymbol{\Theta}_{Mt}|\mathbf{I}_{t-1}) &= Cov(\boldsymbol{\Theta}_{Dt},\, G_t\boldsymbol{\Theta}_{Dt} + \epsilon_t|\mathbf{I}_{t-1}) \\ &= Cov(\boldsymbol{\Theta}_{Dt},\, G_t\boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1}) \\ &= G_t Var(\boldsymbol{\Theta}_{Dt}|\mathbf{I}_{t-1}) \\ &= G_t(H_t C_{t-1} H_t' + W_t) \\ &= G_t H_t C_{t-1} H_t' + G_t W_t \end{aligned} \tag{4.19}$$

That is,

$$R_t = \begin{pmatrix} H_t C_{t-1} H_t' + W_t & G_t H_t C_{t-1} H_t' + G_t W_t \\ H_t C_{t-1} H_t' G_t' + W_t G_t' & G_t H_t C_{t-1} H_t' G_t' + G_t W_t G_t' + \Sigma_t \end{pmatrix}. \tag{4.20}$$

(b) The observational equation in (4.1) and equation in (4.9) lead to

$$(\mathbf{Y}_t|\mathbf{I}_{t-1}) \sim N(\mathbf{f}_t,\, Q_t)$$

where

$$\mathbf{f}_t = E(\mathbf{Y}_t|\mathbf{I}_{t-1}) = E(F_t'\boldsymbol{\Theta}_t|\mathbf{I}_{t-1}) = F_t' E(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1}) = F_t'\mathbf{a}_t \tag{4.21}$$

and

$$\begin{aligned} Q_t &= Var(\mathbf{Y}_t|\mathbf{I}_{t-1}) \\ &= Var(F_t'\boldsymbol{\Theta}_t + \nu_t|\mathbf{I}_{t-1}) \\ &= F_t' Var(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1})F_t + Var(\nu_t|\mathbf{I}_{t-1}) \\ &= F_t' R_t F_t + V_t \end{aligned} \tag{4.22}$$

(c) Suppose the prior distribution of $\boldsymbol{\Theta}_t$ is $p(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1})$ and the likelihood is $p(\mathbf{Y}_t|\boldsymbol{\Theta}_t)$. Then by using Bayes' theorem, the posterior distribution of $\boldsymbol{\Theta}_t$ is given by,

$$p(\boldsymbol{\Theta}_t|\mathbf{I}_t) \propto p(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1})p(\mathbf{Y}_t|\boldsymbol{\Theta}_t) \tag{4.23}$$

Now from the equation (4.11), the prior probability is given by,

$$p(\boldsymbol{\Theta}_t|\mathbf{I}_{t-1}) \propto \exp\left\{-\frac{1}{2}(\boldsymbol{\Theta}_t - \mathbf{a}_t)' R_t^{-1}(\boldsymbol{\Theta}_t - \mathbf{a}_t)\right\} \tag{4.24}$$

and from the observation equation in 4.1, the likelihood is given by,

$$p(\mathbf{Y}_t|\boldsymbol{\Theta}_t) \propto \exp\left\{-\frac{1}{2}(\mathbf{Y}_t - F_t'\boldsymbol{\Theta}_t)' V_t^{-1}(\mathbf{Y}_t - F_t'\boldsymbol{\Theta}_t)\right\} \tag{4.25}$$

By using equations (4.24) and (4.25), we can write the equation (4.23) as,

$$\begin{aligned} p(\boldsymbol{\Theta}_t|\mathbf{I}_t) &\propto \exp\left[-\frac{1}{2}\left\{(\boldsymbol{\Theta}_t - \mathbf{a}_t)' R_t^{-1}(\boldsymbol{\Theta}_t - \mathbf{a}_t) + (\mathbf{Y}_t - F_t'\boldsymbol{\Theta}_t)' V_t^{-1}(\mathbf{Y}_t - F_t'\boldsymbol{\Theta}_t)\right\}\right] \\ &\propto \exp\left[-\frac{1}{2}\left\{\boldsymbol{\Theta}_t(R_t^{-1} - F_t V_t^{-1} F_t')\boldsymbol{\Theta}_t - 2\boldsymbol{\Theta}_t'(R_t^{-1}\mathbf{a}_t + F_t V_t^{-1}\mathbf{Y}_t)\right\}\right] \end{aligned} \tag{4.26}$$

From the theorem statement in equation (4.13), $A_t = R_t F_t Q_t^{-1} = R_t^* F_t V_t^{-1}$ gives

$$\begin{aligned} R_t^* &= R_t - A_t Q_t A_t' \\ &= R_t - R_t F_t A_t' \\ &= R_t(I - F_t A_t') \end{aligned} \tag{4.27}$$

$$\Rightarrow F_t A_t' = I - R_t^{-1} R_t^* \tag{4.28}$$

Now,

$$\begin{aligned}(R_t^{-1} + F_t V_t^{-1} F_t') R_t^* &= R_t^{-1} R_t^* + F_t V_t^{-1} F_t' R_t^* \\ &= R_t^{-1} R_t^* + F_t A_t' \\ &= R_t^{-1} R_t^* + (I - R_t^{-1} R_t^*) \\ &= I\end{aligned} \tag{4.29}$$

where I is an identity matrix , so that

$$R_t^{-1} + F_t V_t^{-1} F_t' = R_t^{*-1} \tag{4.30}$$

with $\mathbf{a}_t^*$ as in the theorem statement in equation (4.13),

$$R_t^{*-1} \mathbf{a}_t^* = R_t^{-1} \mathbf{a}_t - F_t V_t^{-1} \mathbf{Y}_t \tag{4.31}$$

From equation (4.26) we can write,

$$\begin{aligned}p(\boldsymbol{\Theta}_t | \mathbf{I}_t) &\propto \exp\left[-\frac{1}{2}\left\{\boldsymbol{\Theta}_t R_t^{*-1} \boldsymbol{\Theta}_t - 2\boldsymbol{\Theta}_t' R_t^{*-1} \mathbf{a}_t^*\right\}\right] \\ &\propto \exp\left[-\frac{1}{2}\left\{(\boldsymbol{\Theta}_t - \mathbf{a}_t^*)' R_t^{*-1} (\boldsymbol{\Theta}_t - \mathbf{a}_t^*)\right\}\right], \qquad \boldsymbol{\Theta}_t \in \mathbb{R}^n\end{aligned} \tag{4.32}$$

That is,

$$(\boldsymbol{\Theta}_t | \mathbf{I}_t) \sim N(\mathbf{a}_t^*, R_t^*) \tag{4.33}$$

where $\mathbf{a}_t^* = \mathbf{a}_t + A_t \mathbf{e}_t$ and $R_t^* = R_t - A_t Q_t A_t'$,

with $A_t = R_t F_t Q_t^{-1}$, and $\mathbf{e}_t = \mathbf{Y}_t - \mathbf{f}_t$

□

4.5 The Normal Inverse-Wishart HBST Model

Now define another variation of the HBST model that will also be used later in the application in Chapter 5 & 6. This model, called the *normal inverse-Wishart hierarchical Bayesian space-time* (NWHBST) model, attributes a normal prior distribution with parameters $\mathbf{a}_t$ and R_t to $\mathbf{\Theta}_t$ conditional on V_t and $\mathbf{I}_{t-1}$, i.e. $(\mathbf{\Theta}_t | V_t, \mathbf{I}_{t-1}) \sim N[\mathbf{a}_t, R_t]$, and a prior inverse-Wishart distribution with d_{t-1} degrees of freedom and scale matrix Ψ_{t-1} to the uncertain V_t conditional on $\mathbf{I}_{t-1}$, i.e. $(V_t|\mathbf{I}_{t-1}) \sim IW[\Psi_{t-1}, d_{t-1}]$ where IW indicates an inverse-Wishart distribution. The joint distribution of $(\mathbf{\Theta}_t, V_t|\mathbf{I}_{t-1})$ is then a normal-inverse-Wishart model, i.e. $(\mathbf{\Theta}_t, V_t|\mathbf{I}_{t-1}) \sim NIW[\mathbf{a}_t, R_t, \Psi_{t-1}, d_{t-1}]$. See e.g. Theorem 7.7.3 in Anderson (1984).

In this thesis we have adopted an extension of the prior-to-posterior updating approach proposed by Barbosa and Harrison (1992) for the temporal multivariate DLM to include the spatial components of our model. This approach avoids the pragmatic problem usually caused by numerical computations involving matrices of possibly obtaining posterior covariance matrices that are not positive semi-definite and symmetric. Barbosa and Harrison (1992) approach is an approximate conjugate procedure that considers all covariance matrix factorization based on eigenvalue decomposition and (unlike the use of Cholesky decomposition) is invariant to permutations of the elements of $\mathbf{Y}_t$. It has also shown to improve on other filters such as the Robust filter introduced by Masreliez (1975)

and used by West (1982) as an approximation for the marginal posterior distribution of the state parameter vector in a multivariate DLM.

The basic approach is based on the factorisation of V_t with eigenvalue matrix Λ_t and eigenvector matrix D_t such that $V_t = D_t \Lambda_t D_t' = S_t^2$, where $S_t = D_t \Lambda_t^{1/2} D_t'$ is unique, and the setting of the initial point estimate $V_0 = S_0^2$ to act as a reference matrix. Initially, $(\boldsymbol{\Theta}_0 | V_0, \mathbf{I}_0) \sim N[\mathbf{a}_0, R_0]$, where $\mathbf{I}_0$ is any initial information, and $(V_0 | \mathbf{I}_0) \sim IW[\Psi_0, d_0]$ where Ψ_0 of dimension $(k+2)(m+n) \times (k+2)(m+n)$ is positive definite and $d_0 \geq (k+2)(m+n)$. The initial hyperparameters $\mathbf{a}_0$, R_0, Ψ_0 and d_0 are fixed by the user. Thus, at $t = 1$, $(\boldsymbol{\Theta}_1 | V_0, \mathbf{I}_1) \sim N[\mathbf{a}_1, S_0 R_1 S_0]$ and $(V_1 | \mathbf{I}_1) \sim IW[\Psi_1, d_1]$. The full recursive approach is shown in Appendix B.

Note that to run the NWHBST model, apart from specifying the initial prior hyperparameters described above, the user will also need to choose the values for the parameters associated with the HBST model formulation (i.e. G_0, H_0, Σ_0 and W_0) as described in the previous subsections.

The next chapter will describes the HBST model in detail by an application to the radioactivity deposition.

Chapter 5

Application of the HBST Model to Radioactivity Deposition

The previous chapter described the general framework of the HBST model. The chapter describes an application of the HBST model to radioactivity deposition which allows for uncertainty handling and data assimilation of ECOSYS. This chapter starts by describing an overview of the physical models for nuclear emergencies with the different modules in the whole chain where the formulation of HBST model is based on one module. Section 5.2 introduces the HBST model for uncertainty handling and data assimilation associated with the ground deposition of radioactivity after a nuclear accident. In subsection 5.2.1, a graphical representation of the deposition module of ECOSYS is described. Subsection 5.2.2 describes the spatial components of the HBST model and the link between the HBST model and the ECOSYS is established. In subsection 5.2.3, the formulation of HBST model for radioactivity deposition is described. Subsections 5.2.4 and 5.2.5 describe the HBST models with fixed functional spatial covariances and the normal inverse-Wishart HBST model.

5.1 Physical Models for Nuclear Emergencies

The data from a nuclear accidents are not generally available due to the sensitivity of the information and the lack of proper measurement protocols during the emission. Even if available they would probably be useful for the modelling related to the geographic region they come from but not for other regions. Typically, such models have great complexity, lack of linearity and an internal integrity. They can be thought as machines which take a function of outputs from previous modules, vectors of observations and vectors of known covariates as their inputs to produce outputs. The outputs of this module can then be used as one of the components of the inputs of a subsequent module.

A simplified graphical representation of the whole chain of modules for the four mentioned phases (a pre-release phase, release phase, early post-release phase and a late post-release phase) of events in Section 1.1 can be seen in Figure 5.1. Each module (circle node) is associated with a well established engineering or physical/environmental model as we shall see. The square nodes represent the input-output interfaces between modules and the filled circle nodes represent sources of information or data.

The graph gives a general idea of the Markovian nature and the dimension of the problem. Each circle node embeds a particular component model (module) which in its turn can also be represented graphically in a highly structured process. In fact, the *In Plant Status* node is composed of the belief nets mentioned

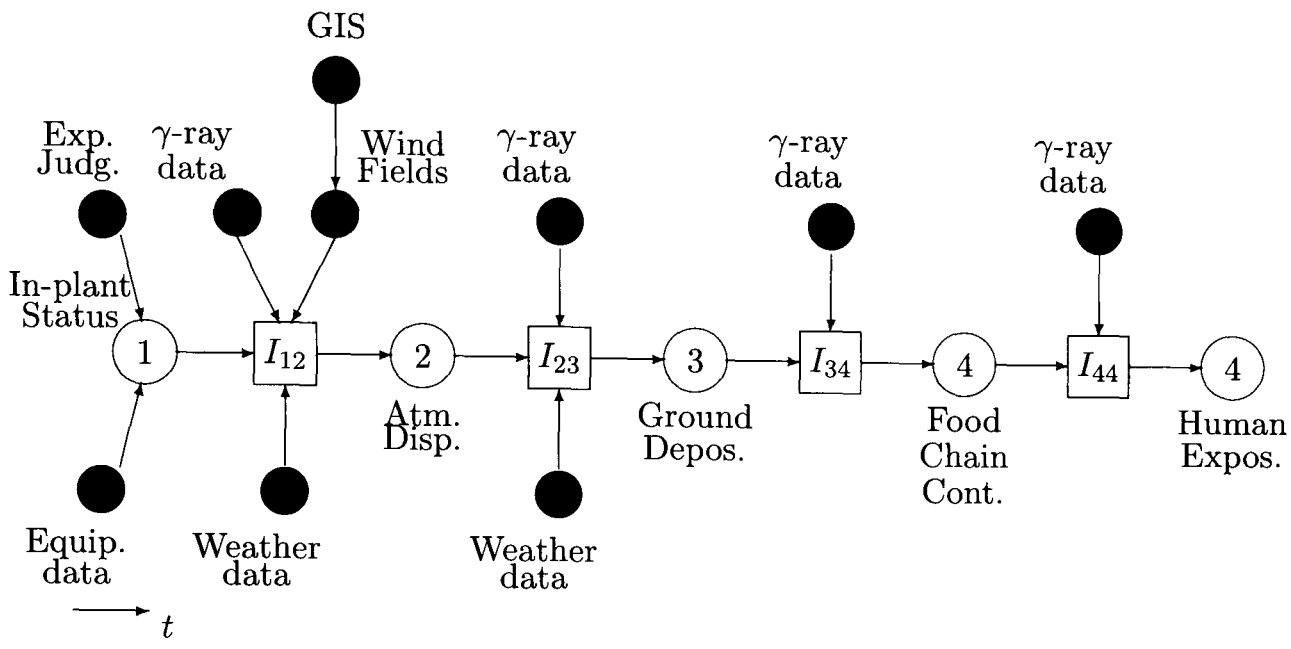


Figure 5.1: Physical models for nuclear emergencies. The following abbreviations are used in this figure: Expert Judgement (Exp. Judg.), Equipment data (Equip. data), Gamma-ray data (γ-ray data), Geographic Information System (GIS), Atmospheric Dispersal (Atm. Disp.), Ground Deposition (Ground Depos.), Food Chain Contamination (Food Chain Cont.) and Human Exposures (Human Expos.).

earlier for source term and release height estimation. Those are based on diverse reactor technologies. They describe the causal associations of possible faults in equipments and circuits in the nuclear plant, and use expert judgements together with equipment data to elicit the conditional probabilities associated with each node (Smedley *et al.*, 1996). The estimated source term and release height can then be used together with measured data (typically there will be a ring of detectors around the nuclear plant with γ-spectrometry and γ-dose rates monitoring capabilities) as input for the following atmospheric dispersal model represented by the node *Atm. Disp.* on Figure 5.1.

Combining the information from the components of the triple (output, observations, covariates) which form the basis of the inputs of the next module over time in a way which: (i) preserves the integrity of each module in the system (i.e. does not interfere with the internal code of that module), (ii) combines observational and other module inputs in a sensible way when these two sources are compatible, and (iii) detects and reacts appropriately when these components contradict one another.

Referring to Figure 5.1, we shall make use of outputs from the node *Atmospheric Dispersal* (Atm. Disp.), such as the radioactivity deposition predictions for a regular spatial grid coming from the K-model, as inputs to the node *Ground deposition* (Ground Depos.) which is a component of the food chain model. Those predictions are treated as data and combined with deposition measurements taken at sites which are usually off-grid locations. In practice, the deposition measurements (or other measurements from which deposition can be calculated, such as air concentrations together with rainfall intensities and gamma dose rates) will not be available for all sites of interest. In fact, there will be a number of fixed measuring stations irregularly and sparsely separated over the region of interest. Some mobile measuring capability should be possible however in small scale.

The development of our integrated hierarchical Bayesian approach is oriented to include the ECOSYS model of Müller and Pröhl (1993) as a particular case. The ECOSYS is a detailed physical-biological model developed to explain the spread of contamination within the food chain with possible consequences on the

human exposure doses. Therefore, it includes the modules *ground deposition*, *food chain contamination* and *human exposure* of Figure 5.1. We adopt a state-space approach similar to the dynamic linear model of West and Harrison (1997) however extended and tailored to our spatio-temporal process. The approach monitors the levels of contamination in space and time, allowing for the process's means and covariances to adapt to data. The updating algorithms given here are obtained in closed form from the normal-inverted-Wishart conjugate analysis. For a description of such models used in a simpler and different context see Le and Zidek (1992).

5.2 HBST Model for Radioactivity Deposition

This section describes the proposed HBST model for radioactivity deposition in detail. The following subsections start by briefly describing the deposition module of ECOSYS as well as the link between the ECOSYS and the HBST model. I have made use of graphical representations in order to facilitate those descriptions. The graphical representations also allow to depict the important conditional independence structure associated with the models. After that, a Gaussian HBST model is defined for radioactivity deposition. To comply with the Gaussian errors assumption, the defined variables and parameters were transformed. In practice, the contamination variables will not be strictly Gaussian since, for example, the contamination values cannot be negative. Also, proportion variables which values belong to the interval (0,1) are obviously non-Gaussian. This

problem is solved by the use of appropriate transformations which allows the treatment of errors as Gaussian as we shall see.

5.2.1 The Deposition Module of ECOSYS

As mentioned in the introduction in Chapter 1, I shall focus on the deposition module of ECOSYS which models multiple-nuclide radioactivity depositions to soil and plants. Figure 5.2 shows a simplified graphical representation of the chain of modules associated with the interface between a long-range Atmospheric Dispersal Model (ADM) and a Food Chain model. In Figure 5.2 each module is represented by a circle node, the input-output (I/O) interface is represented by a square node, the information or data are represented by filled circle nodes and the associations between those are represented by the directed arrows. In the application to radioactivity deposition in this thesis, the HBST model uses K-model (Lauritzen and Mikkelsen, 1999) as the long-range ADM.

In the deposition stage (3) of the accident (described in Section 1.1), ECOSYS calculates for a general site $\mathbf{s}$ with cartesian coordinates $(x, y) \in \mathbb{R}^2$, the total radioactivity $A(i, j, \mathbf{s})$ produced by the deposited isotope type j which is intercepted by the plants type i (e.g. grass, winter wheat, beet, potatoes etc. For further details please see, Müller and Pröhl (1993)). The type j radionuclide deposition to soil at site $\mathbf{s}$ is treated separately for its effects on root uptakes and ressuspension to air. They are otherwise obtained similarly to the plants.

The calculations are performed by decomposing the total deposition of ra-

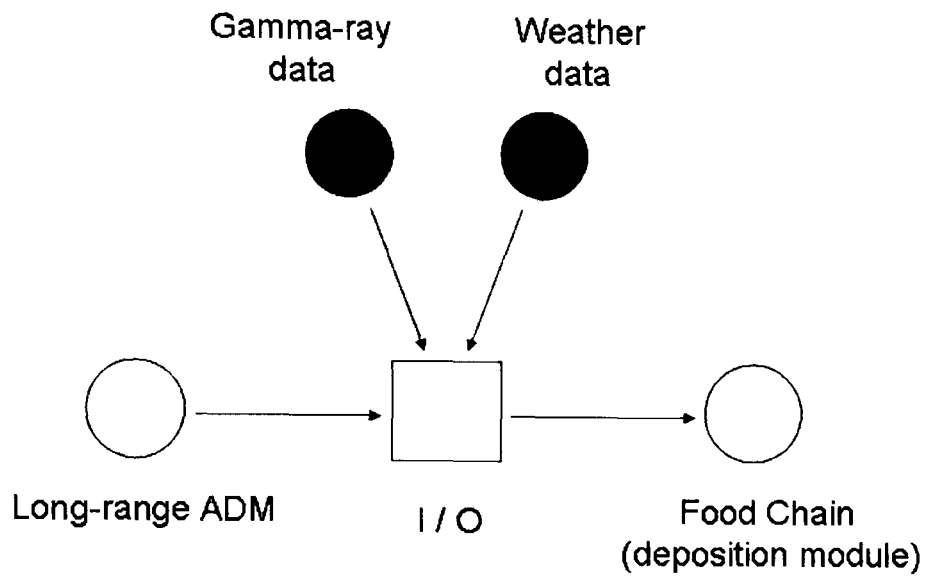


Figure 5.2: A graphical model for the interface between the long-range ADM and the Food Chain deposition module

dioactivity, A, into wet and dry components. The dry component caused by dry deposition counts for the air activity concentration contribution to the contamination while the wet component counts for both air concentration and rainfall contributions. Thus, if $A_{dry}(i, j, \mathbf{s})$ is the component of the total contamination caused by dry depositions of isotope j on plant i at the location $\mathbf{s}$, and $A_{wet}(i, j, \mathbf{s})$ is the wet deposition component of isotope j at $\mathbf{s}$, then

$$A(i, j, \mathbf{s}) = A_{dry}(i, j, \mathbf{s}) + A_{wet}(i, j, \mathbf{s}) \ , \tag{5.1}$$

where $A_{wet}(i, j, \mathbf{s}) = f_w(i, j, \mathbf{s}) A_{wet}(j, \mathbf{s})$ with $f_w(i, j, \mathbf{s})$ called the *interception fraction* of isotope j for plants type i at site $\mathbf{s}$.

The dry contamination component $A_{dry}(i, j, \mathbf{s})$ is obtained from the time-

integrated activity concentration of isotope j in air at $\mathbf{s}$, $\bar{C}_{air}(j, \mathbf{s})$, as

$$A_{dry}(i, j, \mathbf{s}) = v_g(i, j)\bar{C}_{air}(j, \mathbf{s}) \ , \tag{5.2}$$

where

$$v_g(i, j) = v_{g,max}(i, j)\frac{LAI(i)}{LAI_{max}(i)} \tag{5.3}$$

is called the *deposition velocity* of isotope j on the plants type i at $\mathbf{s}$, with $v_{g,max}(i, j)$ being the maximum deposition velocity (i.e. for fully developed foliage), $LAI(i)$ being the *leaf area index* (LAI) for the plants type i, and $LAI_{max}(i)$ the leaf area index (LAI) of plant type i for fully developed foliage.

On the other hand, the interception fraction of wet contamination for plant i is

$$f_w(i, j, \mathbf{s}) = \frac{LAI(i)S(i, j)}{R(\mathbf{s})}[1 - \exp\{-\frac{\ln 2}{3S(i, j)}R(\mathbf{s})\}] \ , \tag{5.4}$$

where $S(i, j)$ is the *retention coefficient* of radionuclides type j in plants type i, and $R(\mathbf{s})$ is the amount of rain of a rainfall event at $\mathbf{s}$.

In an extension of ECOSYS, Bleher and Jacob (1993) developed a washout model in which the wet deposition $A_{wet}(j, \mathbf{s})$ can be obtained from $\bar{C}_{air}$ and $R(\mathbf{s})$ according to the following equation:

$$A_{wet}(j, \mathbf{s}) = h_0\lambda_0(j)\left[\frac{R(\mathbf{s})}{R_0(\mathbf{s})}\right]^{0.8}\bar{C}_{air}(j, \mathbf{s}) \ , \tag{5.5}$$

where h_0 is the assumed effective cloud thickness, $\lambda_0(j)$ is a rate corresponding to a reference rainfall intensity $R_0(\mathbf{s})$, and $R(\mathbf{s})$ is the local rainfall intensity.

In fact, the key variables in obtaining $A(i, j, \mathbf{s})$ in (5.1) are $\bar{C}_{air}(j, \mathbf{s})$ and $R(\mathbf{s})$. Substituting equations (5.2), (5.4) and (5.5) in (5.1) above we obtain

$$A(i, j, \mathbf{s}) = [v_g(i, j) + b(i, j)R^*(i, j, \mathbf{s})]\bar{C}_{air}(j, \mathbf{s}) \tag{5.6}$$

where

$$R^*(i, j, \mathbf{s}) = \frac{[1 - \exp -d(i, j)R(\mathbf{s})]}{[R(\mathbf{s})]^{0.2}}. \tag{5.7}$$

Now, in terms of the statistical modelling of the above module, note that, although non-linear, the first two moments of $A(i, j, \mathbf{s})$ can be easily obtained if independence between $\bar{C}_{air}(j, \mathbf{s})$ and $R^*(i, j, \mathbf{s})$ is assumed. In this case,

(i) $\mathrm{E}[A] = (v_g + b\mathrm{E}[R^*])\mathrm{E}[\bar{C}_{air}]$

(ii) $Var[A] = v_g^2 Var[\bar{C}_{air}] + 2bv_g E[R^*]Var[\bar{C}_{air}] + b^2\{Var[R^*]Var[\bar{C}_{air}] +$
$\mathrm{E}^2[R^*]Var[\bar{C}_{air}] + \mathrm{E}^2[\bar{C}_{air}]Var[R^*]\}$

where $\mathrm{E}[A]$ is the expected value of A and $Var[A] = \mathrm{E}\{(A - \mathrm{E}[A])^2\}$ and the indexes i, k and $\mathbf{s}$ were omitted for simplicity.

Another way of dealing with this non-linearity and the one which the HBST model has adopted, is to assume that $v_g(i, j)$ and $f_w(i, j, \mathbf{s})$ in equations (5.3) and (5.4) above are known and fixed. In practice, this approximation is a good one since the variation of those coefficients is much smaller when compared with the variation of R and $\bar{C}_{air}$. This gives a reasonable linear approximation for the problem and should be suitable for the decision support purposes of our system. Typical or average values of the $LAI(i)$ can be used in fixing $v_g(i, j)$,

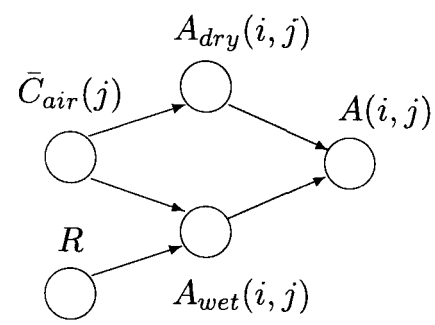


Figure 5.3: A graphical representation of the deposition module in ECOSYS

while estimated values of $R(\mathbf{s})$ could be used in choosing values for $f_w(i,j,\mathbf{s})$. The error introduced by doing this is considered in the modelling of $A_{wet}(i,j,\mathbf{s})$.

Figure 5.3 shows a directed acyclic graph (DAG) depicting the deposition module of ECOSYS. In that graph, a circle (or node) represents a variable and an arrow represent directional (causal) direct effect between variables. It basically depicts the fact that in ECOSYS, the total contamination is directly influenced by both the dry and the wet contaminations which in turn are directly influenced by the total concentration in air. The wet (but not the dry) contamination is also directly influenced by the amount of rainfall. For simplicity, and without loss, we have omitted the location index $\mathbf{s}$ from the variables in the remaining of this section.

Also, note that from the DAG of Figure 5.3, and using the notation of Dawid (1979) –where for three random variables X, Y and Z, $X \amalg Y|Z$ means that X is independent of Y given Z– we can write the following conditional independence statements (see e. g. Smith (1989)):

(a) $\bar{C}_{air}(j) \amalg R$

(b) $A_{dry}(i,j) \amalg A_{wet}(i,j) | \bar{C}_{air}(j)$

(c) $A(i,j) \amalg \bar{C}_{air}(j), R | A_{dry}(i,j), A_{wet}(i,j).$

Conditional independence statements similar to (a)-(c) above defined in the following section are used to simplify the joint probability distribution of the variables associated with the HBST model as we shall see.

5.2.2 The HBST Model's Spatial Components

The HBST model defines a regular grid with n points over the geographic region of interest. Each grid point represents a location for which the long range atmospheric dispersal model will generate a prediction. The network of m fixed measuring stations will be usually located off-grid.

In practice, a measuring station will be able to measure at least one of the following data:

(i) instantaneous air concentrations (c_{air}) and γ-dose rates (γ-Dose) which are informative about the time (and nuclide) aggregated total depositions,

(ii) γ-spectrometry (γ-Spec), informative about the proportions of different radionuclides on the total depositions, and

(iii) rainfall intensity (r), giving information about the proportion of wet deposition relative to the total.

Omitting the time and space indexes (t and $\mathbf{s}$ respectively) for simplicity, the HBST model has three main components Z, P and Q associated with the above described data:

Z: total ground radioactivity deposition,

P_j: proportion of the radionuclide j $(j = 1, \ldots, k)$ in the total ground deposition, and

Q: fraction of wet deposited radioactivity from the total ground deposition.

Those components allow not only the assimilation of the above mentioned measurements but also the uncertainty calculations relative to the nuclide specific wet and dry radioactivity intakes for different types of plants as we shall see.

With the above components introduced, we can draw the *directed acyclic graph* (DAG) of Figure 5.4 which makes a link between our statistical HBST model and a pragmatic version of the deposition module of the food chain ECOSYS model. The main difference between the theoretical ECOSYS described in the previous section and its pragmatic version shown in the DAG below is that the activities in plants $A(i, j, \mathbf{s})$ are obtained from the activities in soil $A(soil, j, \mathbf{s})$. This is due to fact that the model is making use of near-ground measurements of activities as surrogate for measures of deposition in soil. Plant specific activities can be estimated from those with the use of appropriate factors as described by Müller and Pröhl (1993).

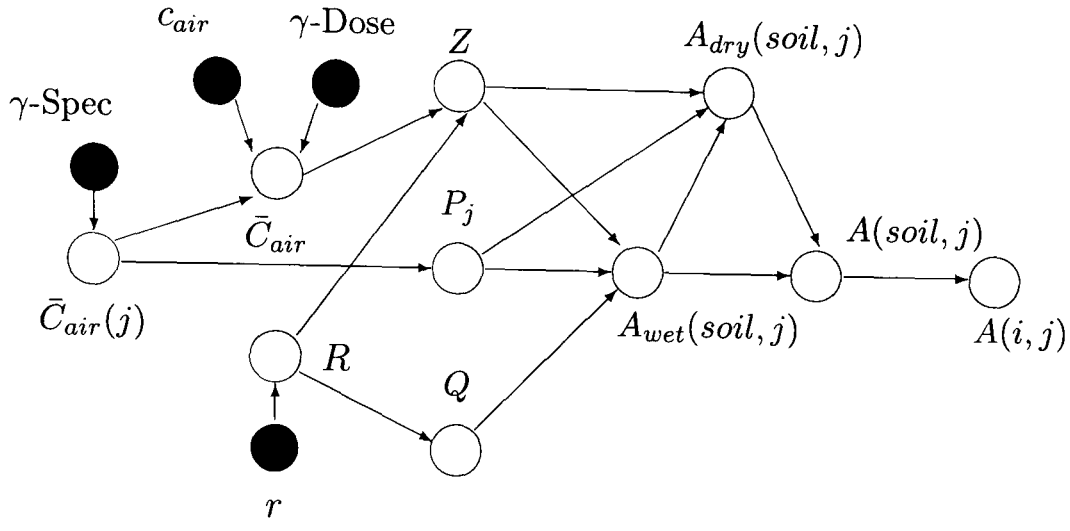


Figure 5.4: A directed acyclic graph linking the HBST model and the ECOSYS model

The graph depicts the logical associations between the variables of the statistical model including the sources of informative data. The filled circle nodes represent sources of information (data) such as γ-Dose rates readings (γ-Dose), γ-spectrometry (γ-Spec), instantaneous air concentrations (c_{air}) and rainfall intensities (r).

Measurements of time-integrated air concentrations, $\bar{C}_{air}(j, \mathbf{s})$, are usually obtained by integrating instantaneous air concentration readings, $C_{air}(j, \mathbf{s}, t)$ in time. They can also be obtained from γ-dose rate readings as described in Bleher and Jacob (1993). However, the isotopic composition in air concentrations needs γ-spectrometry readings to be determined. Readings of γ-ray spectrum in their turn can also give information on measures of instantaneous air concentrations. In fact, γ-dose rates readings give information on total aggregated time-integrated air concentrations $\bar{C}_{air}(\mathbf{s}) = \sum_j \bar{C}_{air}(j, \mathbf{s})$, while γ-spectrum readings give infor-

mation on $\bar{C}_{air}(j, \mathbf{s})$, that is on the components of $\bar{C}_{air}(\mathbf{s})$.

Note that in this graph, the HBST model components Z, $P(j)$ and Q separate the ECOSYS inputs $\bar{C}_{air}$ and R from the activity outputs $A_{dry}(soil, j)$ and $A_{wet}(soil, j)$ (as well as $A(soil, j)$ and $A(i, j)$) such that once those components are known the output variables will have no further (new) information to be gained from the input variables. In fact, the following conditional independencies can be stated from the DAG of Figure 5.4:

(a) $Z \amalg P_j \amalg Q | \bar{C}_{air}, R$ (but $Z \not\amalg P_j \not\amalg Q$ unconditionally);

(b) $A(i, j), A(soil, j) \amalg Z, P_j, Q | A_{wet}(soil, j), A_{dry}(soil, j)$;

(c) $\bar{C}_{air}, R \amalg A_{dry}(soil, j), A_{wet}(soil, j) | Z, P_j, Q$

Those conditional independence statements are used to simplify the joint probability distribution functions of those variables which can be factorized accordingly. Please refer to Smith (1989) for examples of how to read conditional independence statements from directed acyclic graphs as well as to factorize joint probability distribution functions according to such statements.

The nuclide specific activity intake by plant i at a location $\mathbf{s}$, $A(i, j, \mathbf{s})$, can be obtained from $A(soil, j, \mathbf{s})$, since $A_{wet}(j, \mathbf{s}) = A_{wet}(soil, j, \mathbf{s}) = Z_{\mathbf{s}} P_{j,\mathbf{s}} Q_{\mathbf{s}}$ and $A_{dry}(soil, j, \mathbf{s}) = Z_{\mathbf{s}} P_{j,\mathbf{s}}[1 - Q_{\mathbf{s}}]$. In fact, the nuclide specific wet deposited activity at a location $\mathbf{s}$ is assumed to be deposition over soil, that is $A_{wet}(soil, j, \mathbf{s})$ for nuclide j. From that, the wet deposition at $\mathbf{s}$ for grassland is calculated as $A_{wet}(grass, j, \mathbf{s}) = f_w(grass, j, \mathbf{s}) A_{wet}(soil, j, \mathbf{s})$, where $f_w(grass, j, \mathbf{s})$ is the in-

terception fraction of nuclide j for grass at site $\mathbf{s}$. The wet deposition of nuclide j for all other types of plants is calculated in a similar way from $A_{wet}(grass, j, \mathbf{s})$ as $A_{wet}(i, j, \mathbf{s}) = \left[\frac{f_w(i,j,\mathbf{s})}{f_w(grass,j,\mathbf{s})}\right] A_{wet}(grass, j, \mathbf{s})$ for each plant i.

Before defining the HBST model in more detail, I would like to point out that all the variables associated with the components Z, P_j and Q above as well as their related parameters have been transformed. Those transformations were necessary so that the model could comply with the assumption of Gaussian errors. In particular, the HBST model has used logarithmic transformations for the variables associated with total ground deposition as well as neutral to the right (NTTR) transformations of Walker and Muliere (1997) for the variables associated with the proportions components. The NTTR transformations guarantee that the proportion variables add up to unity and can be assumed to be Gaussian.

5.2.3 The HBST Model Formulation

The random variables define below are associated with the components Z, P_j and Q described above. Because vectors of those variables will need to be decomposed into subcomponents grouped according to their origin, we shall use throughout the thesis the subscripts D and M, to indicate the spatial location of data represented, that is, D will represent grid points and M measuring station points.

Let, at a time t $(t = 1, \ldots, \tau)$, $\mathbf{Z}_{Dt} = (Z_{D1t}, \ldots, Z_{Dnt})'$where Z_{Dit} is the

observed total radioactivity deposition at grid point i $(i = 1, \ldots, n)$ at time t, and $\mathbf{Z}_{Mt} = (Z_{M1t}, \ldots, Z_{Mmt})'$ where Z_{Mit} is the observed total radioactivity deposition at measuring station site i $(i = 1, \ldots, m)$ at time t.

Similarly, let $\mathbf{P}_{Dt} = (\mathbf{P}_{D1t}, \ldots, \mathbf{P}_{Dnt})$ with $\mathbf{P}_{Dit} = (P_{Di1t}, \ldots, P_{Dikt})'$ where P_{Dijt} is the proportion (from the total deposition Z_{Dit}) of radionuclide j deposited at the grid point i at time t, and $\mathbf{P}_{Mt} = (\mathbf{P}_{M1t}, \ldots, \mathbf{P}_{Mmt})$ with $\mathbf{P}_{Mit} = (P_{Mi1t}, \ldots, P_{Mikt})'$ where P_{Mijt} is the proportion (from the total deposition Z_{Mit}) of radionuclide j deposited at the off-grid measuring site i at time t.

Also, let $\mathbf{Q}_{Dt} = (Q_{D1t}, \ldots, Q_{Dnt})'$ where Q_{Dit} is the proportion (from the total deposition Z_{Dit}) of wet deposited radioactivity at the grid point i at time t, and $\mathbf{Q}_{Mt} = (Q_{M1t}, \ldots, Q_{Mmt})'$ where Q_{Mit} is the proportion (from the total deposition Z_{Mit}) of wet deposited radioactivity at the grid point i at time t.

Now, let $\mathbf{Y}_{Dt} = (\mathbf{Z}_{Dt}, \mathbf{P}_{Dt}, \mathbf{Q}_{Dt})'$ and $\mathbf{Y}_{Mt} = (\mathbf{Z}_{Mt}, \mathbf{P}_{Mt}, \mathbf{Q}_{Mt})'$. Let $\mathbf{\Theta}_{Dt} = (\theta_{Dt}, \pi_{Dt}, \rho_{Dt})'$, $\mathbf{\Theta}_{Mt} = (\theta_{Mt}, \pi_{Mt}, \rho_{Mt})'$ be the state vectors associated with the observational vectors $\mathbf{Y}_{Dt}$ and $\mathbf{Y}_{Mt}$ respectively, where the components θ_{Dt} and θ_{Mt} are vectors with elements θ_{it} $(i = 1, \ldots, n$ for grid points or $i = 1, \ldots, m$ for measuring station sites) interpreted as the real *unknown* total radioactivity deposition at location i at time t. Similarly, each element π_{ijt} of π_{it} is the real *unknown* proportion of isotope j in the total deposition θ_{it} at location i at time t; and each element ρ_{it} of ρ_{ct} $(c = D$ or $M)$ is the real *unknown* proportion of wet deposited activity from the total deposition at location i at time t. We can now define the HBST model by the following observation, interpolation and time

evolution equations.

- The *observation equation* is

$$\begin{bmatrix} \mathbf{Y}_{Dt} \\ \mathbf{Y}_{Mt} \end{bmatrix} = \begin{bmatrix} \mathbf{\Theta}_{Dt} \\ \mathbf{\Theta}_{Mt} \end{bmatrix} + \begin{bmatrix} \nu_{Dt} \\ \nu_{Mt} \end{bmatrix}, \quad \begin{bmatrix} \nu_{Dt} \\ \nu_{Mt} \end{bmatrix} \sim N(\mathbf{0}, V_t) \tag{5.8}$$

where the observational error terms associated with grid points, $\nu_{Dt} = (\nu_{Z_{Dt}}, \nu_{P_{Dt}}, \nu_{Q_{Dt}})'$, and monitoring stations, $\nu_{Mt} = (\nu_{Z_{Mt}}, \nu_{P_{Mt}}, \nu_{Q_{Mt}})'$, are assumed independent and normally distributed with zero mean vectors and covariance matrix $V_t = diag(V_{Dt}, V_{Mt})$ with $V_{Dt} = diag(V_{Z_{Dt}}, V_{P_{Dt}}, V_{Q_{Dt}})$ and $V_{Mt} = diag(V_{Z_{Mt}}, V_{P_{Mt}}, V_{Q_{Mt}})$ as we shall see. That is, $\nu_{Dt} \sim N[\mathbf{0}, V_{Dt}]$ and $\nu_{Mt} \sim N[\mathbf{0}, V_{Mt}]$.

- The *interpolation equation* is

$$\mathbf{\Theta}_{Mt} = G\mathbf{\Theta}_{Dt} + \epsilon_t \ , \quad \epsilon_t \sim N[\mathbf{0}, \Sigma_t] \tag{5.9}$$

where G is a $(k+2)m \times (k+2)n$ spatial interpolation matrix with fixed elements g_{ij} (defined below) and ϵ_t is the vector interpolation errors assumed normally distributed with zero mean vector and covariance matrix Σ_t.

- The *time evolution equation* is

$$\boldsymbol{\Theta}_{Dt} = H_t \boldsymbol{\Theta}_{Dt-1} + \omega_t \ , \quad \omega_t \sim N[\mathbf{0}, W_t] \tag{5.10}$$

where H_t is a time evolution matrix (described below), and ω_t is the vector of time evolution errors, assumed normally distributed with zero mean vector and covariance matrix W_t.

Now, for $t = 1, 2 \ldots \tau$, from the conditional independence statement (a) in the previous subsection 5.2.2, we can set $V_{Dt} = diag(V_{Z_{Dt}}, V_{P_{Dt}}, V_{Q_{Dt}})$ and $V_{Mt} = diag(V_{Z_{Mt}}, V_{P_{Mt}}, V_{Q_{Mt}})$, where $V_{Z_{Dt}}$, $V_{P_{Dt}}$, $V_{Q_{Dt}}$ and $V_{Z_{Mt}}$, $V_{P_{Mt}}$ $V_{Q_{Mt}}$ are the covariance matrices of the observational errors associated with the HBST model components Z, P_j and Q for grid points and measuring station sites respectively. Note that, the sub-matrices of V_{Dt} are not diagonal themselves, but (due to the Markovian structure of our model) those of V_{Mt} are.

The components of the interpolation error vector $\epsilon_t = (\epsilon_{\theta t}, \epsilon_{\pi t}, \epsilon_{\rho t})'$ are assumed to be mutually independent such that $\Sigma_t = diag(\Sigma_{\theta t}, \Sigma_{\pi t}, \Sigma_{\rho t})$. The covariance matrices $\Sigma_{\theta t}$, $\Sigma_{\pi t}$ and $\Sigma_{\rho t}$ are associated with their respective component parameter (θ, π and ρ respectively). In our application, for each component c ($c = \theta, \pi, \rho$) we have set values of the diagonal elements as

$$\sigma^2_{cit} = \sigma^2_{ct} d^*_i \tag{5.11}$$

where, at each time t, σ^2_{ct} is a common fixed interpolation variance associated with

the component c and $d_i^* = \frac{d_i}{|d_{max}^*|}$, with d_i being the distance of the monitoring station point i $(i = 1, \ldots, m)$ to its nearest grid point and d_{max}^* being the largest distance over all distances from off-grid points to their nearest grid points.

The spatial interpolation matrix is set as $G = diag(G_\theta, G_\pi, G_\rho)$ where G_θ, G_π and G_ρ are spatial interpolation matrices for the components θ_D, π_D and ρ_D respectively. For the application, I have chosen an exponential form for the elements of each submatrices of G, such that

$$g_{ij} = \begin{cases} \alpha_{ij} exp(-\beta_{ij} d_{ij}), & j \in Ne(i) \\ 0, & \text{otherwise.} \end{cases} \tag{5.12}$$

where $Ne(i)$ denotes the set of nearest adjacent grid points to i. α_{ij} is the proportionality constant such that $\sum_j g_{ij} = 1$, $\beta_{ij} \geq 0$ is the rate of exponential decay with the distance d_{ij}. Note that this choice is a fairly reasonable one in applications (like in this thesis) where not only the involved distances are considerably large (as we shall see) but also commonly non-uniform meteorological conditions are largely influential. The adopted exponential form allows the choice of differing spatial decay rates to reflect particular spatial influences at distinct locations in the underlying region for every component of the HBST model. Those could be chosen for example by experts to reflect prevailing terrain structures at those locations.

The deterministic component of the system variables related to their evolution in time is described by the evolution matrix H_t in equation (5.10) which associates $\mathbf{\Theta}_{Dt}$ with $\mathbf{\Theta}_{Dt-1}$. This association is indirectly carried out through

the interpolation equation (5.9). The time evolution matrix H_t is defined as $H_t = diag(H_{\theta_{Dt}}, H_{\pi_{Dt}}, H_{\rho_{Dt}})$ where the component matrices $H_{\theta_{Dt}}$, $H_{\pi_{Dt}}$ and $H_{\rho_{Dt}}$ account for isotopic decays in time of each component of the HBST model.

At each time t, the decay associated with the total deposited activity consists of the aggregation of each deposited radionuclide's decay from the time of the initial deposition. Thus, the sub-matrix $H_{\theta_{Dt}}$ will be diagonal with each diagonal element $h_{\theta_{Dt}} = \sum_j^k h_{\theta_{Djt}}$ independently of spatial location, where

$$h_{\theta_{Djt}} = \exp(-\lambda_j t) \tag{5.13}$$

with $\lambda_j = \frac{ln(2)}{T_{1/2}(j)}$ and $T_{1/2}(j)$ being the half-life of isotope type j, that is, the amount of time required for 50% of the radioactive atoms in a sample to undergo a radioactive (or nuclear) decay.

For practical reasons, we consider the mother-to-daughter decay of isotopes (i.e. the decay of the same element into another type of isotope) as negligible. Thus, the sub-matrix $H_{\pi_{Dt}}$ for the proportions components will be a $k \times k$ diagonal matrix with elements $H_{\pi_{Djt}}$, $j = 1, \ldots k$, where for each j, $H_{\pi_{Djt}}$ is a $n \times n$ diagonal matrix with elements $h_{\theta_{Djt}} = \exp(-\lambda_j t)$ independently of spatial location with λ_j as described above.

The sub-matrix $H_{\rho_{Dt}}$ related to the proportions of wet deposition will also be a diagonal matrix with elements $h_{\rho_{Dt}} = \exp(-\lambda t)$ independent of the spatial location with suitably chosen decay parameter λ.

The initial values, at $t = 0$, of the mean vector $\mathbf{a}_0$ and the covariance matrix

R_0 of the normal distribution of Θ_0 conditional on any initial information must be provided before the model can be used. Also, initial values for the covariance matrices Σ_t and W_t in equations (5.9) and (5.10) respectively, will have to be fixed a priori by the user as we shall see in Chapter 6. Note that, given the measurements of instantaneous air concentrations and amount of rainfall, the diagonal forms of the matrices above can be justified by the conditional independence statement (a) drawn from the DAG of Figure 5.4 in the previous subsection 5.2.2.

In this thesis, we investigate two cases concerning the observational covariance matrix V_t in equation (5.8). In the first case, V_t is assumed *known* and has its values fixed at each time step. In fact, the *known* V_t will have elements with certain fixed functional forms as we shall see in the following subsection. In the second case, V_t is assumed *unknown* and has a prior inverse-Wishart distribution attributed to it as can be seen in Subsection 5.2.5. The first case leads to the conjugate normal model while the second case leads to the conjugate normal-inverse-Wishart model. The prior-to-posterior updating algorithms for the normal and the normal-inverse-Wishart models are described in Appendices A and B respectively. Those algorithms are extensions of the temporal Kalman filter and of Barbosa and Harrison (1992) approximate conjugate method to include the spatial components defined above.

Now define the three variations of the HBST model that will be used later in the application in Chapter 6. Two of those models, defined in the next sub-

section, assume known isotropic spatial covariance matrices that have fixed functional forms. The third model, defined in the subsection after next, assume that the HBST model spatial covariance matrices are unknown and have prior distributions attributed to them.

5.2.4 The HBST Models with Fixed Functional Spatial Covariances for Radioactivity Deposition

Despite the existence of a large number of functional covariance forms available for the elements of valid (i.e. invertible) spatial covariance matrices, I have chosen two particular forms to define two variations of the HBST model. One has an exponential form proposed by Cressie (1993) and the other has a spherical form proposed by Schlather (1999) as described below. The reason behind the choice of those models is pragmatic as they are not only simple to implement but also flexible to use as the spatial decay can be changed with the change of a single parameter.

The *exponential hierarchical Bayesian space-time* (EHBST) model, the HBST model defined by equations (5.8), (5.9) and (5.10) where each component submatrix V_{ct} $(c = Z, P, Q)$ of the observational covariance matrix V_t in equation (5.8) has elements $\sigma_{c_{ijt}}$ $(i = 1, \ldots, m$ and $j = 1, \ldots, n)$ given by

$$\sigma_{c_{ijt}} = \sigma^2_{c_it} exp(-\eta_{c_{ijt}} d_{ij}) \tag{5.14}$$

where $\sigma^2_{c_it}$ is the variance at site i associated with component c at time t, $\eta_{c_{ijt}} > 0$ is a scalar reflecting the strength of the spatial decay in the covariance, and d_{ij} is

the distance between sites i and j. In this model, the spatial correlations between sites i and j decrease exponentially as the distance d_{ij} between them increases which can be seen from the Figure 5.5. The rate of exponential decrease is larger for larger values of the spatial decay parameter $\eta_{c_{ijt}}$ which values must be fixed by the user. Figure 5.5 shows the above exponential function with different spatial decay $\eta_{c_{ijt}} = 0.005$, 0.0189 and 0.050 for all $C = Z$, P, Q. This family often gives a qualitatively reasonably fit to the empirical correlation structure of spatial data.

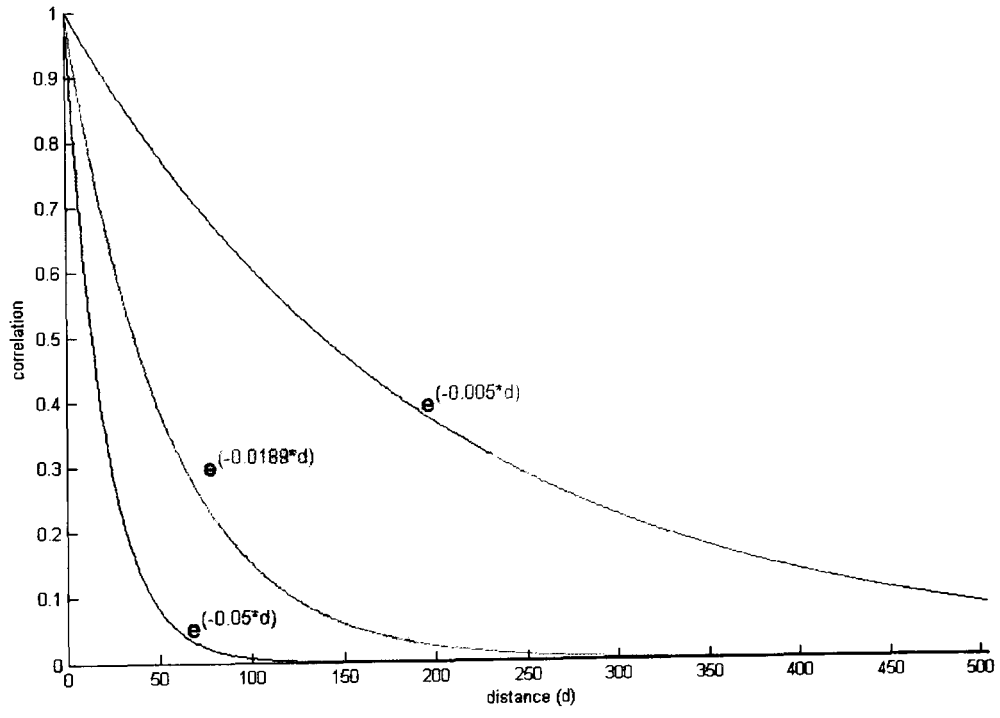


Figure 5.5: Exponential correlation function

Optionally, the *spherical hierarchical Bayesian space-time* (SHBST) model, the HBST model defined by equations (5.8), (5.9) and (5.10) where each component sub-matrix V_{ct} $(c = Z, P, Q)$ of the observational covariance matrix V_t in equation (5.8) has elements $\sigma_{c_{ijt}}$ $(i = 1, \ldots, m$ and $j = 1, \ldots, n)$ given by

$$\sigma_{c_{ij}t} = \sigma^2_{c_it}\left(1 - \frac{3}{2}h_{ij} + \frac{1}{2}h^3_{ij}\right) \tag{5.15}$$

where $\sigma^2_{c_it}$ is the variance at site i associated with component c at time t and $h_{ij} = \frac{d_{ij}}{d_{max}}$ $(0 \leq h_{ij} \leq 1)$ with d_{max} being the largest distance in the set of all distances between grid points. In this model, the spatial correlation function is a polynomial function of order 3 of the distance d_{ij}. A plot of spherical correlation function is shown in Figure 5.6.

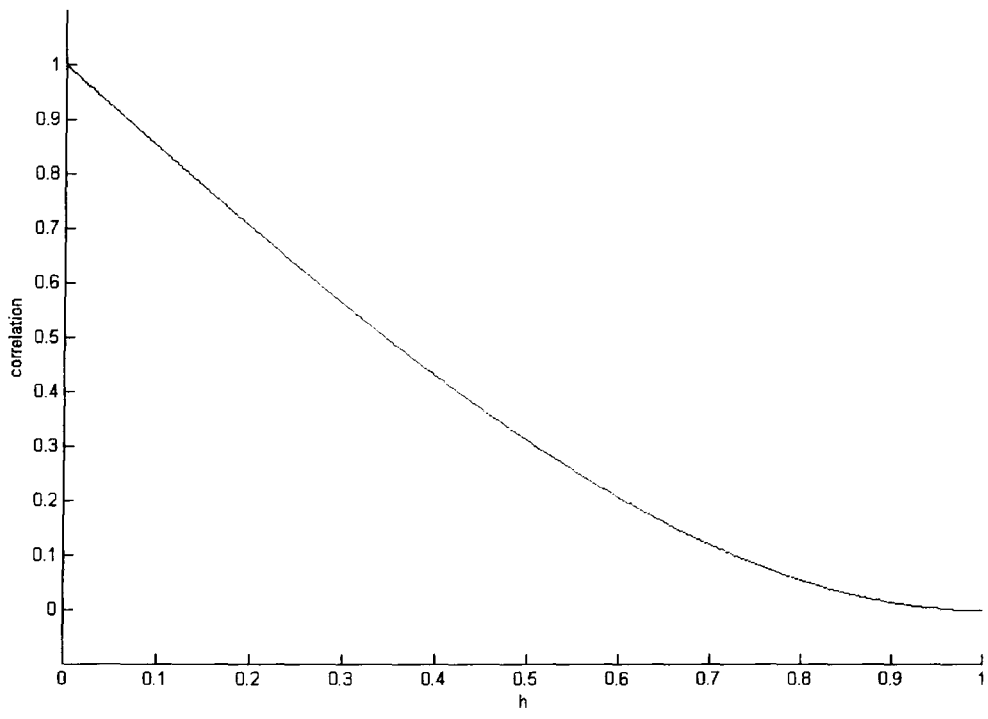


Figure 5.6: Spherical correlation function

Note that compared to the EHBST, the SHBST model is a little less flexible in that the spatial smoothness depends on the distance between the two underlying sites.

In both models, the sub-matrices $V_{Z_{Dt}}$, $V_{P_{Dt}}$ and $V_{Q_{Dt}}$ of V_{Dt} have elements

of the form above. They also have the sub-matrices $V_{Z_{Mt}}$, $V_{P_{Mt}}$ and $V_{Q_{Mt}}$ of V_{Mt} diagonal with fixed variances.

Note that the interpolation matrix G in (5.9) must be specified a priori together with the initial values, at time $t = 0$ (before any observation is taken), of the time evolution matrix H_0 in (5.10) and the observational, the interpolation and the time evolution covariance matrices (i.e. $V_0 = (V_{D0}, V_{M0}), \Sigma_0$ and W_0). Further to those, the mean vector, $\mathbf{a}_0$, and the covariance matrix, R_0, of the Gaussian prior distribution for the initial parameter vector $\mathbf{\Theta}_0 = (\mathbf{\Theta}_{D0}, \mathbf{\Theta}_{M0})'$ must also be specified for both the EHBST and the SHBST models. After those values are fixed, data assimilation and the Bayesian sequential prior-to-posterior updating for those models are performed with the use of the extended Kalman filter algorithm described in Appendix A.

5.2.5 The NWHBST Model for Radioactivity Deposition

A third variation of the HBST model, called NWHBST model has already been defined in Chapter 4. This subsection describes the application of the NWHBST model to radioactivity deposition. The NWHBST model formulation for radioactivity deposition is same as described the NWHBST model in Section 4.5 of Chapter 4. Please see Section 4.5 for details of the NWHBST model formulation and the full recursive approach shown in Appendix B.

Now that the models we will be applying in Section 6.2 of the next chapter are defined. In the next chapter we will proceed to describe the data object of

that application, that is, the radioactive near-ground concentrations in Bavaria caused by the Chernobyl accident in April 1986.

Chapter 6

The Chernobyl Contamination of Bavaria

The proposed HBST model was applied to real data on the 1986 Chernobyl's accident contamination of Bavaria in southern Germany. In fact, three versions of the HBST model combined predictions at fixed regular grid points from the long range atmospheric K-model (Lauritzen and Mikkelsen, 1999) mentioned in the previous chapter with measured (off-grid) near-ground contamination in Bavaria to produce estimations (actual and predictive) of the contamination profile.

This chapter describes the results from the HBST model for radioactivity deposition. Section 6.1 describes the radioactivity data required for HBST model. More specifically, in Subsection 6.1.1, the real data of Chernobyl's near-ground radioactive deposition in Bavaria is described and in Subsection 6.1.2, the predicted data from a long range atmospheric dispersal model, called K-model, is described. Section 6.2 summaries the posterior and predictive results of the HBST model and also the validation of those results with the measurements. Section 6.3 describes the HBST model sensitivity. In subsections 6.3.1 and 6.3.2, we checked

the model sensitivity for the variation in the input parameters of the K-model and the variation in spatial decay of the spatial covariance function for the fixed functional HBST model. The cross-validation of the HBST model is described in Section 6.4 where the whole data set split into two parts, one is treated as the training set and the other is as the validation set.

6.1 The Chernobyl Contamination of Bavaria

This section describes in detail both the measured, posterior and the predicted near-ground radioactive contamination of Bavaria caused by the Chernobyl disaster. The measurements were taken from 13 fixed measuring station sites, while the predictions were those from a long range atmospheric dispersion model, the K-model, for the $4,096$ points of a 64×64 grid over the Bavarian region.

According to the World Nuclear Association (WNA, 2008), Chernobyl's nuclear accident on the 26 April 1986 released an estimated 14 EBq (i.e. 14×10^{18} becquerels) of radioactivity to the atmosphere during a period of at least nine days. Half of the released radioactivity was of biologically inert noble gases. Most of the released material was deposited as dust and debris near Chernobyl, but lighter material was carried by wind over The Ukraine, Belarus, Russia, Scandinavia and Europe, including Bavaria, in southern Germany, as we shall see in this section.

The OECD's Nuclear Energy Agency (NEA, 2002) stated that initial assessment of releases made by Soviet scientists based on sampling from material

deposited in the former Soviet Union, estimated that all of the core inventory of biologically inert noble gases was released, and between 10% and 20% of iodine-131 (I131), caesium-134 (Cs-134) and caesium-137 (Cs-137). Iodine and caesium elements are of particular concern as they are known to be hazardous to human health. The I131 isotope has a short radioactive half-life (eight days), but can be transferred to humans relatively rapidly from the air and through the consumption of contaminated milk and leafy vegetables. Caesium isotopes on the other hand have relatively longer half-lives (Cs134 has a half-life of 2 years while Cs137 has 30 years). Many other less hazardous radioisotopes such as Ruthenium-103 (Ru103), Ruthenium-106 (Ru106) and Tellurium-132 (Te132) amongst others, were also released during the accident.

Weather conditions at the time of the accident were such that only after the third day of the accident prevailing winds were in the western European direction. In fact, it was only from the 29 April 1986 that the measuring station sites in Bavaria started to read higher than normal measurements of radioactivity as we shall see.

The reader may refer to the WNA (2008) or to the NEA (2002) for a more in-depth historical description of the accident and its consequences.

6.1.1 The Radioactivity Data in Bavaria

Daily data of Chernobyl's near-ground radioactivity air concentrations in Bavaria were obtained from the Bavarian regional office for environmental protection

(Böllmann *et al.*, 1987) for the period between 29 April and 10 May 1986 inclusive. The data consisted of γ-spectrometry (γ-spec) readings converted to measures of nuclide specific (i) instantaneous air concentrations, in units of Becquerel per cubic meter (Bq/m^3), and (ii) wet deposited activities, in Becquerel per litre (in Bq/l). Those readings were taken from 13 fixed measuring station sites covering the Bavarian region with an area of approximately $70,549\ Km^2$ as shown in Figure 6.1. Figure 6.1 also shows the 64×64 regular grid covering the whole Bavaria that was used by the HBST models in this application. The distance between consecutive points in the grid is 8 Km.

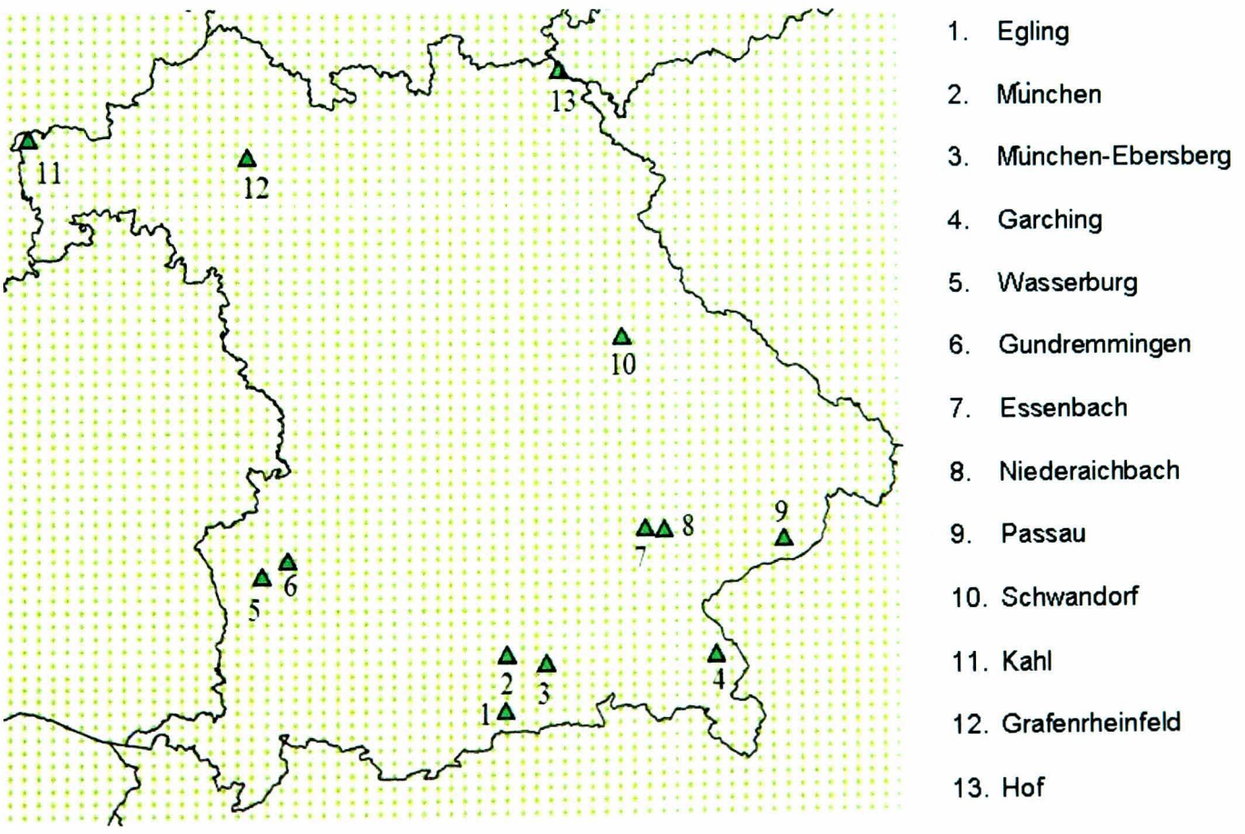


Figure 6.1: The map of Bavaria with the 13 fixed monitoring stations (indicated by triangles with individual numbers on the map and corresponding names on the right hand side table) and the 4096 points of the 64×64 grid over the region. The distance between consecutive grid points is 8 Km.

The data correspond to six isotopes: Ru103, Ru106, Te132, I131, Cs134 and Cs137. Note that measurements of instantaneous air concentrations could be obtained for each of those six isotopes in all 13 monitoring stations, while measurements of wet depositions were only available for 5 (viz. Hof, Kahl, München, Passau and Wasserburg) of the 13 stations.

Note that, with the exception of a few, most stations are rather scattered over the region and quite distant from one another. Most stations (9 of the 13) are located in the south of the region with the majority of those (7 of the 9) situated towards the Bavarian south-east. There is a single station in the center-east Bavaria (Schwandorf) which seems quite spatially isolated distancing by 120 Km from its nearest station (Essenbach). The northern Bavaria region has only three stations (Kahl, Grafenrheinfeld and Hof) which are located far apart from each other. The Hof station also seems quite isolated from the remaining stations. There is a large area in the center-west region with no measuring station whatsoever. Further to those, there are about three spatial clusters of stations (München and Ebersberg, Gundremmingen and Wasserburg, and, Essenbach and Niederaichbach), all in the south, which are relatively closely located. The two nearest stations are Essenbach and Niederaichbach with a distance of approximately 10 Km separating them. The important point to note here is that there are relatively few and scattered monitoring stations to cover the region of interest.

The plots in Figure 6.2 show the time series of measurements of nuclide specific instantaneous air concentrations for 8 measuring stations in the period from the

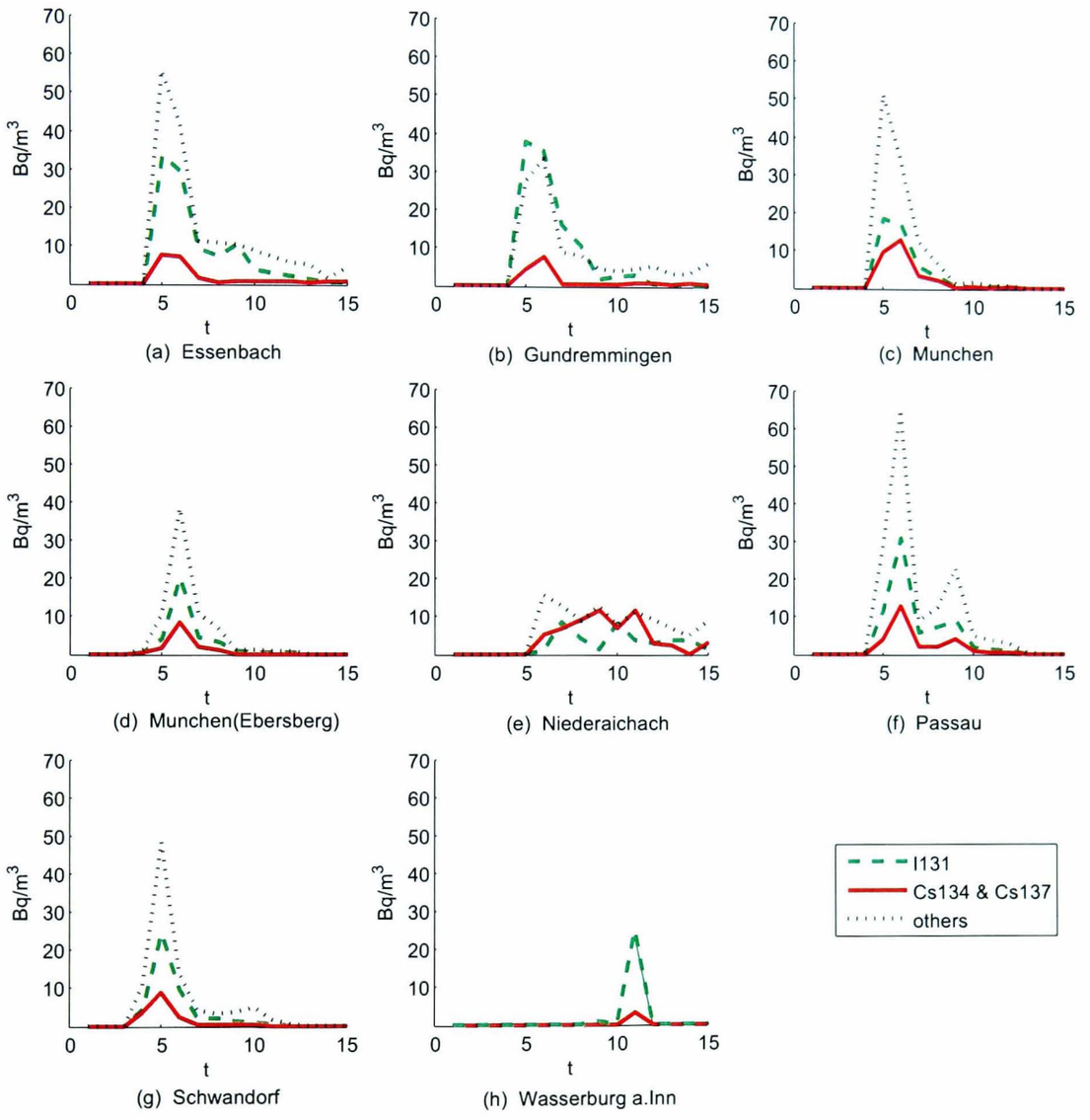


Figure 6.2: Daily average instantaneous near-ground air concentrations in Bq/m^3 of I131 (dashed lines), Caesium (Cs134 and Cs137, solid lines) and the other (Ru103, Ru106 and Te132, dotted lines) radioisotopes from the 26 April ($t = 1$) to the 10 May 1986 ($t = 15$) at (a) Essenbach, (b) Gundremmingen, (c) München, (d) Ebersberg, (e) Niederaichbach, (f) Passau, (g) Schwardorf, and (h) Wasserburg

26 April to 10 May 1986. Data for the other 5 stations are not shown because near-ground concentrations at those sites were only marginally larger than the normal levels. In each plot, the time (horizontal) axis indicate the day number (we adopted $t = 1$ for 26 April, ..., $t = 15$ for 10 May 1986) and the vertical axis indicate the nuclide specific daily average air concentrations (in Bq/m^3). Also, for simplicity, the isotopes were grouped into three categories according to their levels of hazard to human health and half-lives. Thus, the less hazardous isotopes (Ru103, Ru106 and Te132) were labeled as *others* while the most hazardous were split into the shorter lived *Iodine* (I131) with a half-life of 8 days, and the longer lived *Caesium* (Cs134 and Cs137) with half-lives of 2 and 30 years respectively. In fact, from the radiological point of view, I131 and Cs137 are the most important radionuclides to consider for their potentially damaging health effects. The I131 radioisotope accumulates in the thyroid gland and in high doses can cause thyroid cancer while Cs137 mainly concentrates in bones and can cause leukaemia and cancer. See e.g. Castronovo (1999) and Watari *et al.* (1988) for more detailed descriptions of potential health effects of I131 and Cs137 intakes.

We can see from Figure 6.2 that the concentration levels for all the detected radioisotopes at most stations increased dramatically from their normal background levels of $0.012 Bq/m^3$ on average before and in the early days of the accident (from $t = 1$ to $t = 4$) to much higher levels (in the range from 10 to 50 Bq/m^3) in the period from $t = 5$ to $t = 10$, i.e. from the fourth to the ninth day after the reactor's explosion.

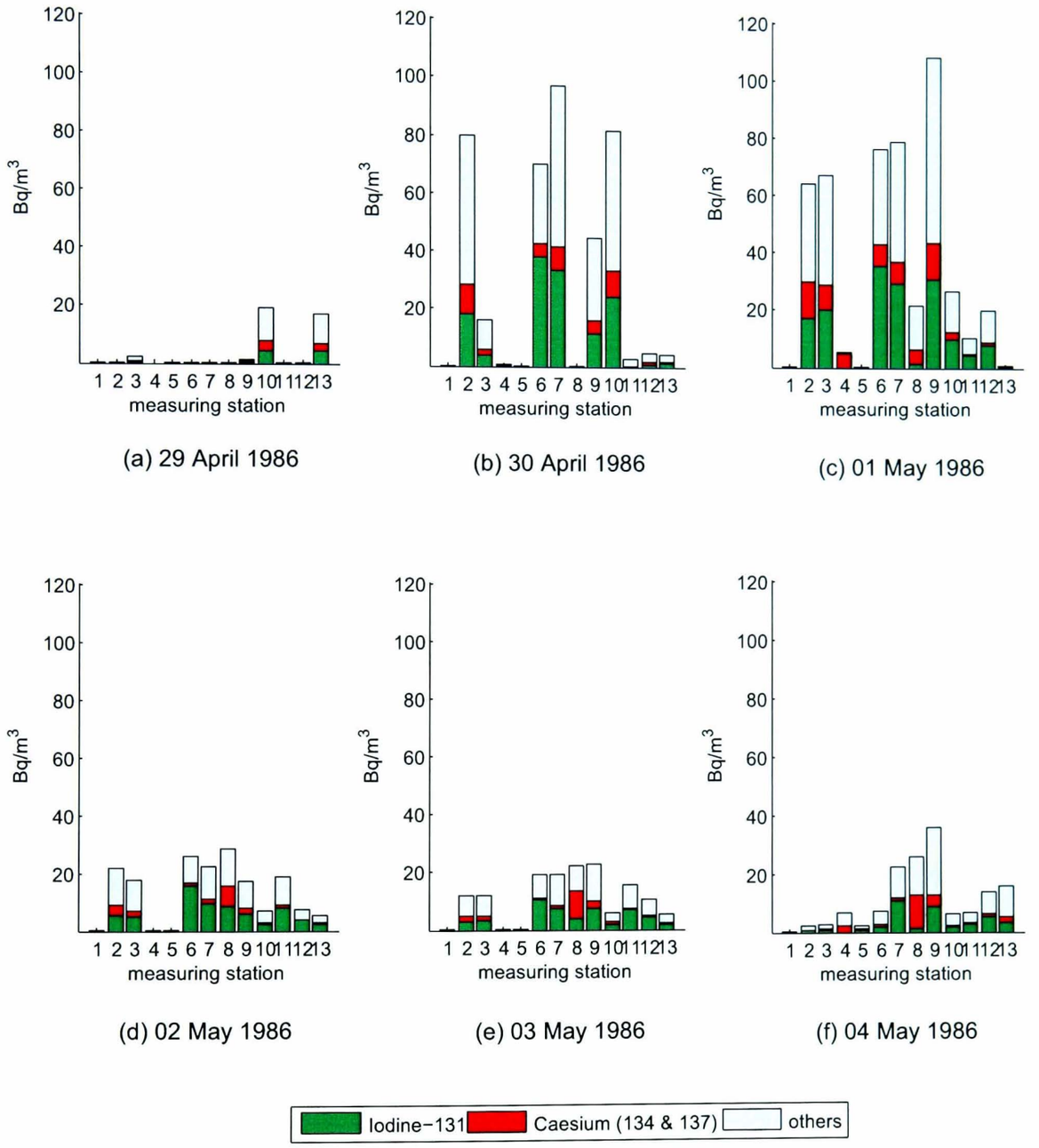


Figure 6.3: Near-ground daily average air concentrations of Iodine (I131), Caesium (Cs134 and Cs137) and other less hazardous isotopes (Ru103, Ru106 and Te132) from the 29 April ($t = 4$) to the 4 May 1986 ($t = 9$) at each of the 13 measuring stations in Bavaria: 1. Egling, 2. München, 3. München-Ebersberg, 4. Garching, 5. Wasserburg, 6. Gundremmingen, 7. Essenbach, 8. Niederaichbach, 9. Passau, 10. Schwandorf, 11. Kahl, 12. Grafenrheinfeld and 13. Hof.

Figure 6.3 shows for each day from the 29th of April to the 4th of May 1986, the nuclide specific near-ground radioactivity air concentrations measured at each of the 13 measuring stations for the three isotope groups: iodine (I131), caesium (Cs134 and Cs137) and others (Ru103, Ru106 and Te132). Notice that both on the 30th April 1986 and on the 1st May 1986, the concentration levels of all isotopes were not only much higher than at any other day but also were particularly so at the measuring stations mainly located in the south-eastern and central Bavaria (Essenbach, Gundremmingen, München, Passau and Schwandorf). In general, the more hazardous isotopes (iodine and caesium) seemed to constitute about 50% of the total concentrations with the proportions of iodine being almost always much larger than the caesium. Also, note that the total daily concentrations had been at their natural average levels (i.e. 0.012 Bq/m^3 or less) up to the 30 April 1986 when the total levels increased enormously to approximately 36.859 Bq/m^3 in average (that is, about 3071 times the natural average level). This is consistent with the fact that the prevailing wind direction at Chernobyl towards Western Europe occurred only after the 29 April as in the weather map of Figure 6.4. Before that, from the 26 April, the wind direction mainly pointed north towards Belarus and Russia.

The meteorological conditions at Chernobyl together with the varying characteristics of the release led to a complex pattern of atmospheric transport and ground deposition both within the Soviet Union and in other countries. However, the rise in concentration levels on the fourth and fifth days after the accident is

consistent with the fact, reported by the NEA (2002), that 'initially the wind was blowing in a Northwesterly direction and was responsible for much of the deposition in Scandinavia, the Netherlands and Belgium and Great Britain. Later the plume shifted to the South and much of Central Europe, as well as the Northern Mediterranean and the Balkans, received some deposition, the actual severity of which depended on the height of the plume, wind speed and direction, terrain features and the amount of rainfall that occurred during the passage of the plume.'

6.1.2 The Long Range Atmospheric Dispersal K-Model

The K-model, proposed by Lauritzen and Mikkelsen (1999), is a long range atmospheric transport model of radionuclides that has been chosen in this thesis, for its simplicity and good performance in real (Chernobyl accident) and simulated applications (Lauritzen *et al.*, 2006), to adopt in the application in this thesis. It produces estimates of instantaneous daily average radioactive concentrations for the grid points over Bavaria (transported from Chernobyl).

The K-model is a first order approximation of a two-dimensional advection-diffusion process which relates turbulent fluxes to the gradient of the mean concentration in the long-range dispersion of a pollutant. In fact, the K-model describes the isotropic diffusion of the near-ground mean radioactivity concentration c_{air} for a radionuclide j at a location $\mathbf{s}$ at time t, in Bq/m^3, by the following equation:

$$c_{air}(j,\mathbf{s},t) = \frac{\Gamma\Delta_t}{2\pi K} e^{-\frac{d(\mathbf{s})v_{t-h}}{2K}} K_0(d^*_{j\mathbf{s}}) \tag{6.1}$$

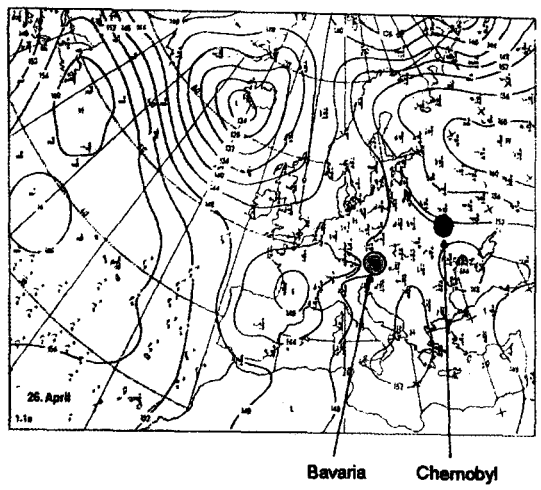
where $d(\mathbf{s})$ is the distance, in Km, from location $\mathbf{s}$ to the source of release: Δ_t is the total radioactivity, in Bq, released from the nuclear reactor at time t; K is the large-scale eddy diffusivity coefficient, in m^2/s; Γ is the ensemble total activity removal rate (or 'effective scavenging coefficient'), in s^{-1}; $v_{t-h(\mathbf{s})}$ is the average wind speed at time $t-h$ at the source of release ($h(\mathbf{s})$ is the length of time the radioactive plume takes to reach the location $\mathbf{s}$ where the concentration is being estimated), in m/s; and, $K_0(d^*_{j\mathbf{s}})$ is the zero-order modified Bessel Function at $d^*_{j\mathbf{s}} = (d(\mathbf{s})/2K)\sqrt{4K\Gamma + \nu_j^2}$, where ν_j is the deposition velocity, in m/s, for isotope j (assumed constant for each radionuclide). Please refer to Lauritzen and Mikkelsen (1999) for more details.

The parameters K and Γ account for a suitably chosen ensemble of long-range weather patterns and are estimated with the use of numerical weather prediction model data. The application in this thesis has adopted the values $K = 10^6 m^2 s^{-1}$ and $\Gamma = (1/864000)s^{-1}$ (corresponding to $10days)^{-1}$) adopted by Lauritzen and Mikkelsen (1999) for the Chernobyl accident. We also determined, for each day, the average wind speed by projecting the prevailing wind vector at Chernobyl, given by Böllmann *et al.* (1987), in the southwest line towards Bavaria. Figure 6.4 shows the weather maps in Europe on (a) the 26 April, (b) 27 April, (c) 28 April, (d) 29 April, (e) 30 April, and on (f) the 01 May 1986. Notice that the prevailing wind directions at Chernobyl in period from 26 to

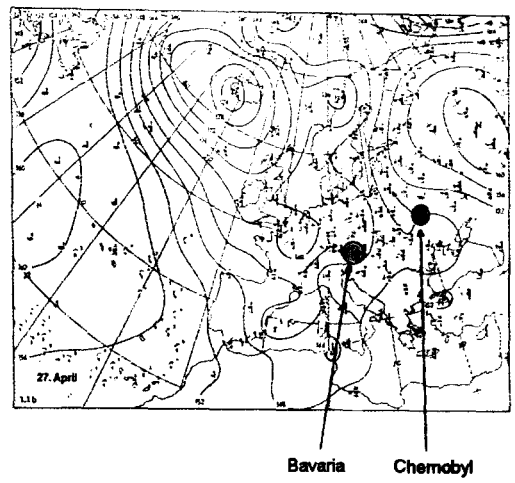
29 of April 1986 were mainly pointing northwest, north and northeast towards Belarus, Scandinavia and Russia. From the 30 April, those changed direction and started to point southwest and south towards Western Europe. However, despite the wind blowing north-northwesterly initially, the average wind speeds during that period were highest so that the projected speed towards Bavaria peaked at $6.21m/s$ ($22.37\ Km/h$) on the 26 April and fluctuated at around $4.23m/s$ ($15.24\ Km/h$) during the following 5 days. From the 1 May, the projected wind speed dropped steadily from $4.29m/s$ ($15.43\ Km/h$) to its minimum value of $1.18m/s$ ($4.26\ Km/h$) on 6 May 1986. It peaked again on the 7 May to reach $5.90m/s$ ($21.25\ Km/h$). As the releases from Chernobyl were taking about three days to reach Bavaria, we have set $h(\mathbf{s}) = 3$ in equation (6.1).

Also we adopted the estimated daily releases for I131, Caesium (134 & 137) and the other isotopes (Ru103, Ru106, Te132) from the nuclide specific estimated total releases provided by the NEA (2002) report and from the daily release fractions given in ApSimon *et al.* (1989). Those values were used to obtain Δ_t in equation (6.1) above.

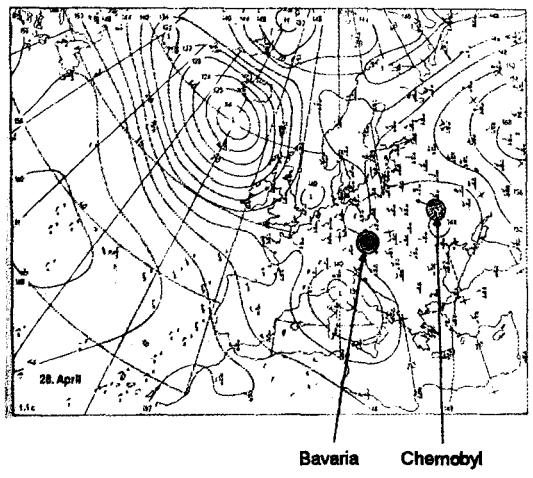
In fact, according to the NEA (2002), Chernobyl's accident estimated total radioactive release of I131, Caesium and the other radioisotopes (Ru103, Ru106 and Te132) together during the whole accident was in excess of 3.29 EBq. Table 6.1 shows the estimates of radionuclide released (Devell *et al.*, 1995) during the Chernobyl accident with the core inventory and half-life of the isotopes. In the Table 6.1, the estimated released activity is shown in Peta-becquerels (PBq)



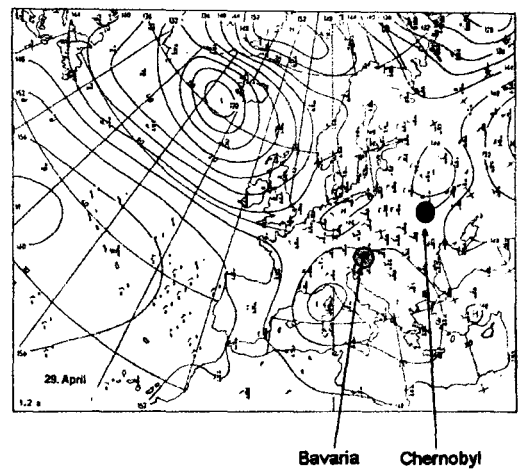
(a)



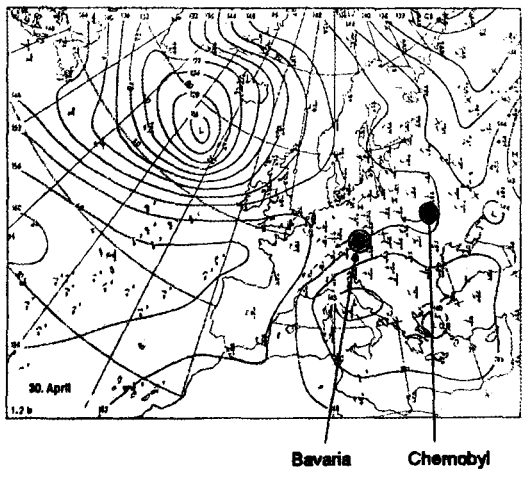
(b)



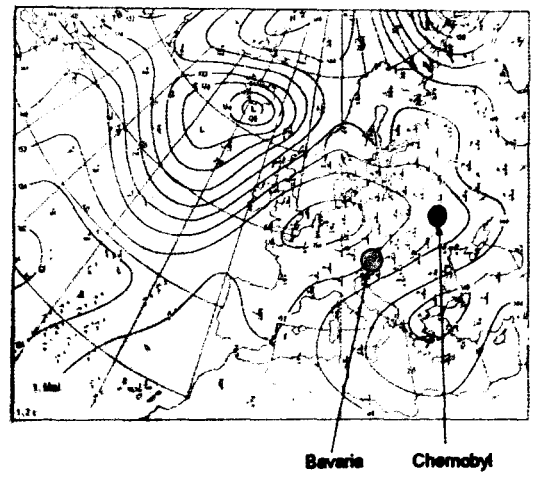
(c)



(d)



(e)



(f)

Figure 6.4: European weather charts with wind vectors at Chernobyl on the (a) 26 April, (b) 27 April, (c) 28 April, (c) 29 April, (c) 30 April, and on the (d) 01 May 1986.

($1PBq = 10^{15}Bq$) and the half-lives are shown in hours (h), days (d) and years (y) respectively. Also, ApSimon *et al.* (1989) estimated the daily percentages from the total release as 23.90%, 7.97%, 6.97%, 4.98%, 3.98%, 3.98%, 7.97%, 9.96%, 13.95% and 15.94% for each respective consecutive day in the period from 26 April to 5 May 1986, as well as 0.20% for both 6 and 7 May 1986 as shown in Figure 6.5.

Table 6.1: Estimates of radionuclide releases during the Chernobyl accident

Core inventory on 26 April 1986			Total release during the accident	
Nuclide	Half-life	Activity (PBq)	Percent of inventory	Activity (PBq)
Ru103	39.6 d	4800	>3.5	>168
Ru106	1.0 y	2100	>3.5	>73
I131	8.0 d	3200	50-60	~1760
Te132	78.0 h	2700	25-60	~1150
Cs134	2.0 y	180	20-40	~54
Cs137	30.0 y	280	20-40	~85

The shortest distance between Chernobyl's border and its nearest point in the grid over Bavaria is 1875 Km. The distances used in equation (6.1) for other points in the grid considered the 8 Km distances between consecutive grid points. Those distances together with the values of the variables described above where used to determine the daily near-ground air concentrations estimated by the K-model at the grid points.

Because the K-model is a long range dispersal model and does not account for local atmospheric variables, the distribution of its estimated daily concentrations was rather smooth in space. This smoothness had a detrimental effect in the HBST models which produced confidence intervals for the estimated concen-

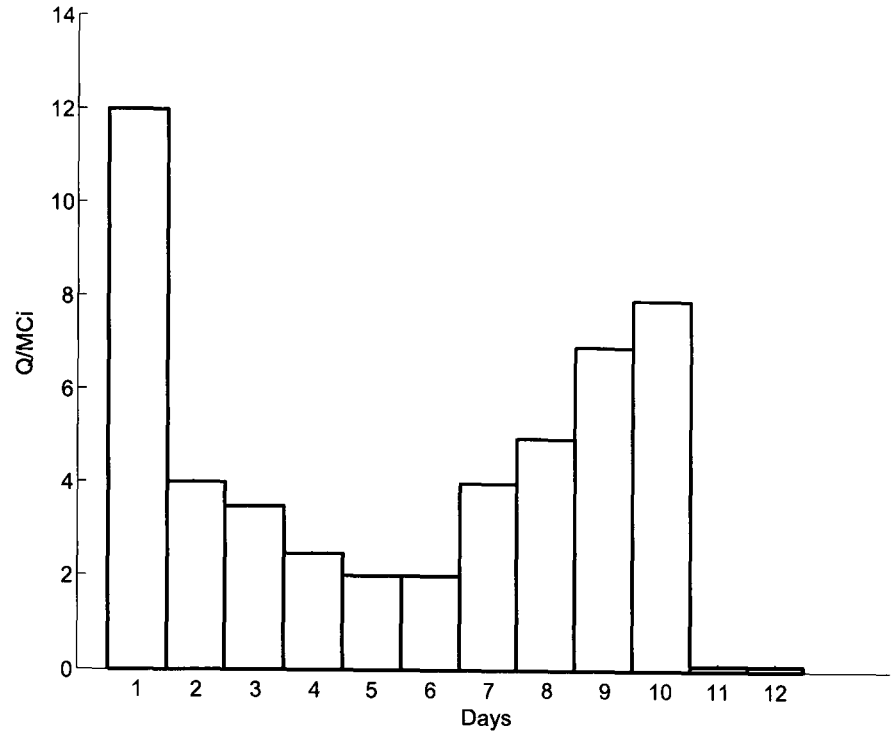


Figure 6.5: Daily variation in release rates

trations which were larger near measuring sites than at grid points. In order to produce more realistic estimates which also produced a less smooth spatial distribution of the estimated concentrations, we have made use of precipitation data (amount of rainfall) at 24 locations in Bavaria during the underlying period of time. Daily average washout measures for 24 locations in Bavaria, obtained from tutiempo.net (2008), were converted into proportions of total washouts, spatially interpolated to grid points and applied to the values estimated by the original K-model to produce what we call *corrected* K-model estimates. Those corrected daily radioactivity concentrations were then used by the HBST models.

The predicted grid-point concentrations, interpolated to the 13 off-grid measuring sites, showed to be consistent with the measured values at those sites. In

fact, those interpolated predictions gave values which were significantly closer to the observed concentrations than interpolations from the K-model's estimations alone (without the washout correction). The corrected predictions produced a *square-Root of the spatial Average of the Mean Square Errors* (RAMSE) of 10.32 Bq/m^3 that was 63.30% lower than the K-model alone. The RAMSE is calculated by

$$RAMSE = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\frac{1}{\tau}\sum_{t=1}^{\tau} e_{it}^2\right)} \tag{6.2}$$

where m is the total number of measuring stations, τ is the total number of release days and e_{it} is the estimation (or prediction) error by the underlying model at site i and time t. This measure of performance is adopted by the various applied models in the following section. In this section, it was used to measure the estimation performance of the corrected K-model when interpolated to the measuring station sites. So, the error term e_{it} in the above equation was the difference between the daily measurement at each site and the interpolated value by the corrected K-Model estimation. When used to measure the predictive performances of the HBST models, the error e_{it} is the difference between the observed value (i.e. either an off grid measurement or a grid point value estimated by the corrected K-model) and the underlying HBST model one-step-ahead median forecast for site i and day t (as defined in subsections A.4 and B.4 in Appendices A and B respectively).

The corrected K-model predictions of near-ground total radioactivity concen-

trations on the 29th of April were very low for the whole region, having varied between 3 and 4 Bq/m^3 that day. However, on the 30th of April, the corrected predictions increased to the range between 30 and 32 Bq/m^3 almost everywhere in the region (except for some rainfall gauging sites where higher precipitation measures implied slightly larger predicted values). The predictions reached their peaks on the 1st of May when they varied between 35 and 38 Bq/m^3 almost everywhere with the exception of some rainfall gauging sites in southern Bavaria where predictions reached between 40 and 43 Bq/m^3. Table 6.2 show the summary statistics (aggregated for all the radionuclides together) of the grid values from the K-model before and after compiling the local rainfall. From Table 6.2, we note that the standard deviations increased enormously (approximately 27.84 times on an average) after compiling the local rainfall with the K-model predictions between 29th April and 10th May 1986. In particular, the increased standard deviation varied between 1.06 times (on 29th April) and 137.00 times (on 9th May). That is, the local rainfall measures have contributed to a less homogeneous and more realistic spatial distribution of the predicted concentrations.

Figure 6.6 shows the contour plot of the predicted Bavarian concentrations for the 1st of May. The dots with individual numbers on the map represent the 24 rainfall gauging sites used to correct the K-Model's predictions. Note that the plot is fairly consistent not only with the K-model's predictions which concentrations decay exponentially but slowly in space as the distance from the origin

Table 6.2: Summary statistics for the grid values for all the radionuclides from K-model in Bavaria (all the measurements are in Bq/m^3)

Day	Without local rainfall				With local rainfall			
	Mean	Min	Max	SD	Mean	Min	Max	SD
29 April 1986	3.06	2.90	3.21	0.0784	3.06	2.90	3.71	0.0833
30 April 1986	31.01	30.90	31.14	0.0606	31.01	30.63	40.90	0.3948
01 May 1986	36.53	35.21	37.93	0.6873	36.53	35.21	43.24	0.7616
02 May 1986	13.27	13.17	13.38	0.0531	13.27	13.00	21.82	0.2756
03 May 1986	11.02	10.91	11.13	0.0561	11.02	10.72	18.00	0.2301
04 May 1986	11.72	11.55	11.89	0.0860	11.72	11.40	16.34	0.1824
05 May 1986	7.43	7.38	7.49	0.0270	7.43	7.31	11.12	0.1248
06 May 1986	9.16	9.14	9.19	0.0138	9.16	9.10	10.63	0.0733
07 May 1986	3.99	3.98	3.99	0.0012	3.99	3.96	4.81	0.0354
08 May 1986	1.95	1.95	1.95	0.0001	1.95	1.93	2.17	0.0136
09 May 1986	1.79	1.79	1.79	0.0001	1.79	1.78	2.08	0.0137
10 May 1986	1.65	1.62	1.69	0.0194	1.65	1.60	2.27	0.0291

(Chernobyl) increases, but also with the corrections for rainfall which introduce changes to the regularity of the contour lines as well as to the regions near the gauging sites in the southern region where predicted concentrations were much larger than nearby regions. The local rainfall measures have indeed contributed to a less homogeneous and more realistic spatial distribution of the predicted concentrations. The actual rainfall measures were large in the southern region (average 0.66mm) and there were no rainfall in the central and northern region on 1st May 1986. In particular, München and Muhldorf (near Essenbach) had the largest rainfall (1.02mm) on 1st May 1986. In general, between 29th April and 1st May 1986 the rainfall measures were large in the southern region only and no rainfall from center to north. On the 2nd of May there were no rainfall measurements in almost all the stations except Muhldorf and from 3rd May to 5th May there were no rain in the southern and central region except the northern region (average 0.40mm). After that, the average rainfall increases on 6th May (1.64mm) and it reaches its peak on 8th May (5.48mm) and again decreases to 0.45mm on an average between 9th and 10th May 1986. Note that, the rainfall measures come from almost all the region in Bavaria during 6th May to 10th May 1986.

On the 2nd of May the predictions dropped to the range between 13 and 14 Bq/m^3 almost everywhere and then from the 3rd to the 6th of May fell to the range between 8 and 12 Bq/m^3. From the 7th to the 10th of May, the corrected predictions were very low again ranging between 1.5 and 4 Bq/m^3

almost everywhere. Note that those predictions were all consistent with both the accident's dispersal behavior and the measurements described in the previous Section.

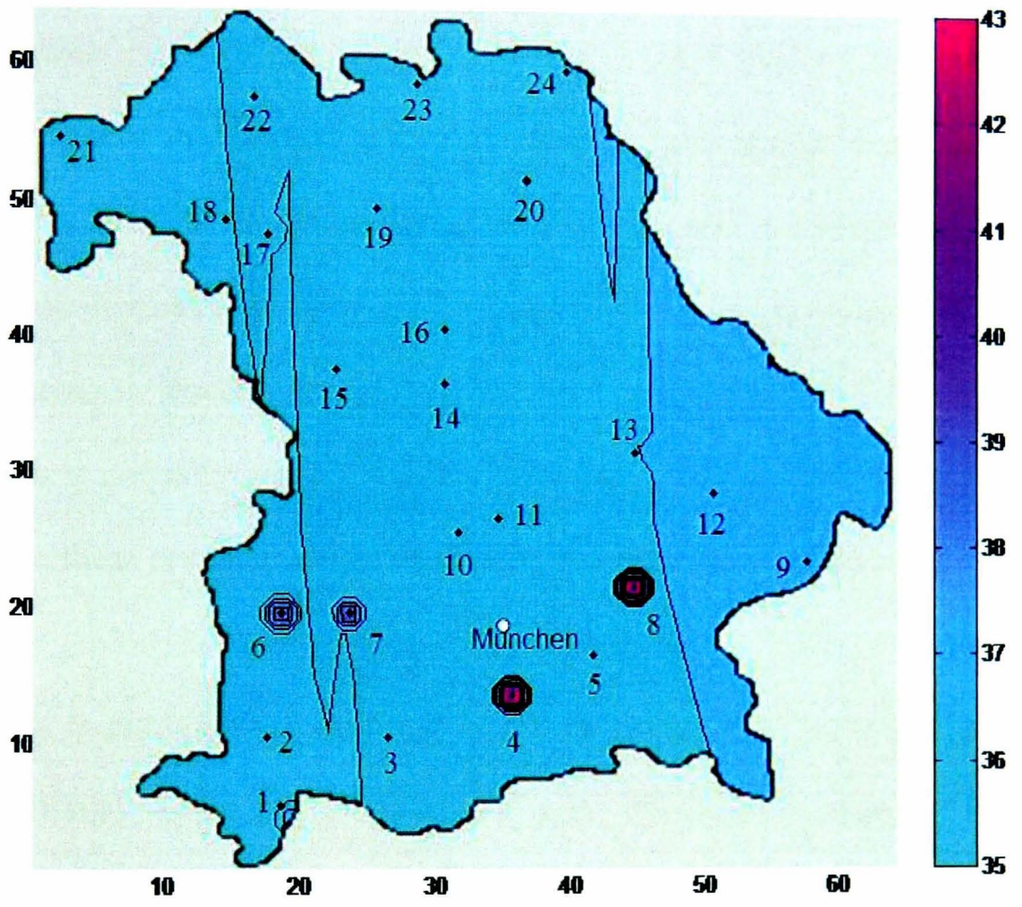


Figure 6.6: Contour plot of the K-model's corrected predictions of total near-ground radioactivity concentrations for the 1st May 1986. The dots with individual numbers represent the locations of the 24 rainfall gauging stations which measurements were used to correct the K-model's predictions.

We can now proceed to describe the results of applying the three variations of the HBST model (viz. the EHBST, the SHBST and the NWHBST models described in Chapter 5) using the rainfall corrected predictions from the K-model described above, to the Chernobyl's near-ground radioactivity concentrations in

Bavaria described in Subsection 6.1.1.

6.2 The HBST Models Applied to the Bavarian Radioactivity Concentration Data

This section describes the main results of the application of the three HBST models, defined in Chapter 5, to the radioactive near-ground concentrations in Bavaria caused by the Chernobyl accident, presented in Section 6.1. Nuclide specific predictions of near-ground concentrations are made for some of the main urban areas in Bavaria for each of the three isotope groups (I131, Cs134-137 and others) defined earlier. Those are compared with observed (or interpolated) values at those urban areas on the days the concentrations were highest to assess the best predictive model. One of the aims of the analyses presented in this section is to illustrate the potential usefulness of the HBST models, in particular the NWHBST, were they to be adopted at the time of the accident.

Subsection 6.2.1 introduces the values chosen for the input parameters as well as the hyperparameters of the prior distributions associated with the HBST models. In subsection 6.2.2 and 6.2.3, the summary of posterior results and its validation against measurements are described. Subsection 6.2.4 shows a comparative analysis of the three models in terms of their predictive performances for both one-day-ahead (we call 'short' term) as well as for three-days-ahead ('medium' term) forecasting of total near-ground radioactivity concentrations in Bavaria. Contour plots of one-day-ahead predictions of total near-ground ra-

dioactivity concentrations and their related uncertainties are also shown for a spatial comparison of the predictive outcomes by all three models. In subsection 6.2.5, bar charts of the nuclide specific concentrations at the measuring sites predicted by the best predictive model, the NWHBST, are shown and compared with the similar bar charts of the measurements displayed in Figure 6.3. Further to those, one-day-ahead forecasts of nuclide specific concentrations are produced by the NWHBST model for five of the largest cities in Bavaria for the 30th April and the 1st May 1986 (the days of highest observed concentrations in Bavaria).

6.2.1 The Input Parameters of the HBST Models

The input parameters required for the observation equation (5.8) are $V_{Dt} = diag(V_{Z_{Dt}}, V_{P_{Dt}}, V_{Q_{Dt}})$ and $V_{Mt} = diag(V_{Z_{Mt}}, V_{P_{Mt}}, V_{Q_{Mt}})$. For the EHBST and the SHBST models, the variances $\sigma^2_{c_i}$ of V_{Dt} in (5.14) were estimated from the available data to give $\sigma^2_{Z_{Di}} = 1.0385$, $\sigma^2_{P_{1Di}} = 0.0011$, $\sigma^2_{P_{2Di}} = 0.0001$, $\sigma^2_{P_{3Di}} = 0.0005$ for iodine, caesium and others (subscripts 1, 2 and 3 respectively) and $\sigma^2_{Q_{Di}} = 1.8026$ for all sites i. Similarly, each component of V_{Mt} was estimated from the data to give $\sigma^2_{Z_{Mi}} = 9.5283$, $\sigma^2_{P_{1Mi}} = 1.9329$, $\sigma^2_{P_{2Mi}} = 2.7459$, $\sigma^2_{P_{3Mi}} = 6.9204$ and $\sigma^2_{Q_{Mi}} = 17.4571$ for all sites i. The values of $\eta_{c_{ijt}}$ for the EHBST model in equation (5.14) were estimated from the observed data and for simplicity adopted independently of spatial location as $\eta_{c_{ijt}} = 0.0189$ for all c, i, j, t. The distance d_{max} in equation (5.15) was set as $d_{max} = 712.76$Km. So, the covariance matrix $V_{Z_{Dt}}$ for the EHBST and the SHBST models have elements $\sigma_{Z_{Dijt}} =$

$1.0385\exp(-0.0189d_{ij})$ and $\sigma_{Z_{Dijt}} = 1.0385(1 - 0.0021d_{ij} + 1.3808 \times 10^{-9}d_{ij}^3)$ respectively.

The input parameters in equation (5.9) are $G = diag(G_\theta, G_\pi, G_\rho)$ and $\Sigma_t = diag(\Sigma_{\theta t}, \Sigma_{\pi t}, \Sigma_{\rho t})$. For simplicity, we have adopted for all three models (independently of spatial location) a common value for the spatial decay parameters $\beta_{ij} = \beta = 0.2$ for all i, j. Also, for simplicity, the values of α_{ij}, for the elements $g_{ij} = \alpha_{ij}\exp\{-\beta_{ij}d_{ij}\}$ of all the submatrices of G, were set to be α_i for all $j \in Ne(i)$. Those values of α_i for $i = 1, 2 \ldots 13$ were estimated from the observed data based on the assumption that $\sum_j g_{ij} = 1$ for all $j \in Ne(i)$ to give $0.73, 0.98, 0.63, 0.78, 1.34, 0.53, 0.77, 0.77, 0.76, 0.76, 0.73, 0.74$ and 0.76 for $i = 1, \ldots, 13$ respectively.

Each submatrix Σ_{ct} $(c = \theta, \pi, \rho)$ of Σ_t is also diagonal with elements σ_{cit}^2 as in equation (5.11). A common value $\sigma^2 = 8.46 \times 10^{-7}$ (estimated from data) was chosen for σ_{cit}^2 for all c, i and t. For each i, $d_i^* = \frac{d_i}{d_{max}^*}$ with $d_{max}^* = 10.21$Km.

The input parameters in equation (5.10) are $H_t = diag(H_{\theta_{Dt}}, H_{\pi_{Dt}}, H_{\rho_{Dt}})$ and the covariance matrix $W_t = diag(W_{\theta t}, W_{\pi t}, W_{\rho t})$. All the submatrices of H_t are diagonal and for each component we adopted the same diagonal values independently of spatial location. Each diagonal element of $H_{\theta_{Dt}}$ is $h_{\theta_{Djt}} = h_{\theta_D} = \sum_{i=1}^{6}\exp(-\lambda_i t)$ independently of each spatial location for all the isotope types j $(j = 1, ..., 6)$. The diagonal elements of $H_{\pi_{Dt}}$ are $\exp(-\lambda_1 t)$, $\sum_{j=2}^{3}\exp(-\lambda_j t)$ and $\sum_{j=4}^{6}\exp(-\lambda_j t)$ for each group of radionuclides (e.g. iodine, caesium and

others) and the each diagonal element of $H_{\rho_{Dt}}$ is $h_{\rho_{Dt}} = \exp(-\lambda t)$ with $\lambda = 0.05$ independently of each spatial locations. The value of λ is suitably fixed by the user. As the wet deposition mainly depends on rainfall, so we have chosen a quite fast exponential decay so that the effect of rainfall covers only the neighboring region of the place where rainfall happens. The diagonal elements of all submatrices of H_t are of exponential form as described in equation (5.13). The parameters $\lambda_j = \frac{ln(2)}{T_{1/2}(j)}$ were determined from the half-life $T_{1/2}(j)$ of each isotope type j $(j = 1, ..., 6)$. The half-life values of the radioisotopes I-131, Cs-134, Cs-137, Ru-103, Ru-106 and Te-132 are 8.04, 753.72, 11033.95, 39.27, 372.3 and 3.204 days respectively. Note the relatively short half-lives of I-131 and Te-132 compared with those of Cs-134, Cs-137, Ru-103 and Ru-106.

For the EHBST and the SHBST models, the elements of each submatrix W_{ct} of W_t described in subsection 5.2.3 are w_{cijt}, which have fixed exponential form similar to $\sigma_{c_{ijt}}$ described in equation (5.14).

Finally, for the data assimilation process by the EHBST and the SHBST models, we fixed the initial values of those models' hyperparameters, i.e. $\mathbf{m}_0$ and C_0. These values are suitably chosen by the users. One possible choice is the non-informative prior distribution of $\mathbf{\Theta}_0$. On the other hand, when the data represents the continuation of a previously observed series, then the time origin $t = 0$ is just being an arbitrary label. In such cases, the initial prior is viewed as sufficiently summarising the information from the past, $\mathbf{\Theta}_0$ having the concrete interpretation of the final state vector for the historical data. In this application,

we fixed the vector $\mathbf{m}_0$ as the average of all the grid point values independently of each spatial locations. The elements of the initial covariance matrix C_0 is chosen as an exponential form similar to $\sigma_{c_{ijt}}$ at $t = 0$ described in equation (5.14).

The NWHBST model adopted the algorithm described in Appendix B. The prior parameters at time $t = 1$ for the NWHBST model are $\mathbf{a}_1$, R_1, Ψ_0 and d_0. To obtain the values of $\mathbf{a}_1$ and R_1, we need the initial mean vector $\mathbf{m}_0$ and the initial variance-covariance matrix C_0 as described above. Ψ_0 is the initial scale matrix for the inverse-Wishart distribution. The implementation of the Barbosa and Harrison updating algorithm described in Appendix B requires the computation of two matrix square roots in B.4 of Appendix B. These matrix square roots are implemented in two different ways, using respectively two different matrix factorisation techniques: the Cholesky decomposition method and the eigen-value decomposition method described in Appendix C. For Cholesky decomposition method Ψ_0 has to be a positive definite matrix and for eigen-value decomposition method Ψ_0 should be symmetric square matrix. The choice of Ψ_0 depends on the users. In our application we choose Ψ_0 from a exponential family which guaranteed the positive definiteness of the matrix Ψ_0. We used the elements of the initial scale matrix Ψ_0 is chosen as an exponential form similar to $\sigma_{c_{ijt}}$ at $t = 0$ described in equation (5.14) and d_0 is the initial degrees of freedom with value in the present application being $d_0 = (m + n) \times (k + 2) = 20545$, where $m = 13$, $n = (64 \times 64)$ and $k = 3$ (all were defined earlier). S_0 is the square-root of the initial observational matrix V_0, where $V_0 = \Psi_0/d_0$.

With all the input parameters and hyperparameters specified, the Chernobyl data assimilation was carried out by the EHBST and the SHBST model through the spatio-temporal Kalman filter described in Appendix A and by the NWHBST model through the updating algorithm described in Appendix B.

6.2.2 Summary of the Posterior Results of the HBST Models

We can now proceed to describe the posterior results of applying the three variations of the HBST model (viz. the EHBST, the SHBST and the NWHBST models described in Chapter 5) using the rainfall corrected predictions from the K-model described in Subsection 6.1.2, to the Chernobyl's near-ground radioactivity concentrations in Bavaria described in Subsection 6.1.1.

With all the input parameters specified in the previous subsection, the HBST models can perform the data assimilation through the spatio-temporal Kalman filter described in Appendix A and the Barbosa and Harrison (1992) updating algorithm described in Appendix B. Calculated posterior results are summarized in Table 6.3.

To compare the model performances, we calculated the RAMSE for all the three HBST models, as described in equation (6.5). We also calculated the spatial Average Mean Absolute Deviation (AMAD) for the three HBST models described in equation (6.3). The AMAD is calculated by

Table 6.3: Posterior medians (Med.) and 95% credible intervals for the parameters of the original variables of total near-ground concentrations of the EHBST, the SHBST and the NWHBST models calculated on the 01 May 1986 at the 13 measuring stations (Egling (M_1), München (M_2), München-Ebersberg (M_3), Garching (M_4), Wasserburg (M_5), Gundremmingen (M_6), Essenbach (M_7), Niederaichbach (M_8), Passau (M_9), Schwandorf (M_{10}), Kahl (M_{11}), Grafenrheinfeld (M_{12}) and Hof (M_{13})).

θ	EHBST			SHBST			NWHBST		
	Median	2.5%	97.5%	Median	2.5%	97.5%	Median	2.5%	97.5%
θ_{M_1}	28.33	6.37	126.03	34.37	7.89	149.69	18.04	17.62	18.47
θ_{M_2}	69.46	15.86	304.24	58.05	13.59	247.94	39.54	38.62	40.47
θ_{M_3}	62.52	14.43	270.85	42.73	11.26	162.12	36.81	35.96	37.67
θ_{M_4}	25.20	7.60	83.64	20.42	6.75	61.79	17.05	16.73	17.38
θ_{M_5}	27.65	6.24	122.62	31.77	7.93	127.34	17.76	17.35	18.19
θ_{M_6}	95.76	49.99	183.41	96.46	61.57	197.42	108.89	97.27	120.97
θ_{M_7}	60.20	26.40	137.30	123.10	20.55	737.49	91.02	88.40	93.72
θ_{M_8}	53.45	12.52	228.16	37.13	10.40	132.60	35.30	34.50	36.13
θ_{M_9}	53.03	12.53	224.38	35.69	10.06	126.60	34.46	33.68	35.26
$\theta_{M_{10}}$	53.31	12.10	234.97	34.31	9.38	125.52	31.25	30.52	32.00
$\theta_{M_{11}}$	39.94	9.70	164.45	33.44	9.77	114.41	25.24	24.68	25.81
$\theta_{M_{12}}$	50.06	11.70	214.17	38.11	10.05	144.54	31.10	30.39	31.83
$\theta_{M_{13}}$	12.26	4.59	32.70	11.01	4.74	25.57	8.87	8.73	9.01

$$AMAD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{\tau}\sum_{t=1}^{\tau}|e_{it}|\right)} \tag{6.3}$$

where $N = (m+n)$ is the total number of grid points and the measuring station sites, τ is the total number of release days and $|e_{it}|$ is the absolute value of the estimation error by the underlying model at site i and time t.

Table 6.4 summaries the RAMSE and the AMAD values for the HBST models. According to Table 6.4, the overall performance of the NWHBST model is better than EHBST and SHBST models. That is, the HBST model with the learning from data on covariance fits better to the data compared to the known and fixed functional spatial covariance models.

Table 6.4: The HBST models performances

	EHBST	SHBST	NWHBST
RAMSE	0.5876	2.3835	0.2258
AMAD	0.4399	0.9501	0.1077

In all the contour plots in this thesis, low concentrations are in *cyan* and high concentrations are in *pink*. Intermediate values are displayed in various intensities of pink, the darker the intensity the larger the concentration values.

The contour plots (a) to (c) in Figure 6.7 display the posterior mean of the total near ground radioactivity air concentrations in Bavaria on 01 May 1986 by the three HBST models (viz. the EHBST, SHBST and the NWHBST models). The contour maps are based on the posterior distribution of θ_{Dt} given the whole data $(\mathbf{Y}_{Dt}, \mathbf{Y}_{Mt})$, where θ_{Dt} is the unknown parameter for total near ground

radioactivity deposition at the grid points D at time t.

We noticed that in general the posterior means varied from $100Bq/m^3$ to $110Bq/m^3$ for the EHBST model, where as for the NWHBST model it varied from $30Bq/m^3$ to $40Bq/m^3$ at the region near München, Passau and Regensburg, i.e. the north-east and the south-east region of Bavaria. On the other hand, for the SHBST model the posterior mean varied from $45Bq/m^3$ to $55Bq/m^3$ near München and near Passau it reached $70Bq/m^3$ to $80Bq/m^3$. Comparing those posterior means with the measured values in some measuring stations located in that region shown in Figure 6.3, we notice that in general on 1st May 1986 the NWHBST model produced better results (i.e. close to the measured values) compare to the EHBST and SHBST models. We also noticed that the two local maximum of the total near ground deposition appeared near München and Essenbach (For the exact location of Essenbach please see the Figure 6.1) where the posterior means of the total near ground radioactivity air concentrations reached $120Bq/m^3$ to $130Bq/m^3$ for the EHBST model, $70Bq/m^3$ to $80Bq/m^3$ for the SHBST model, and $40Bq/m^3$ to $50Bq/m^3$ for the NWHBST model. This is because, in those region the rainfall measurements were larger and more accurate than the other region on 01 May 1986.

The contour plots (a) to (c) in Figure 6.8 show the corresponding uncertainty maps for the mean near ground radioactive deposition in Bavaria on 01 May 1986. By incorporating this information, we can greatly improve on the inference provided by the mean field alone. Note that the posterior uncertainty maps for

all the three HBST models are higher at those region where the posterior mean deposition is also higher. Also note that the posterior uncertainties are lower in the north-west region of Bavaria. In the north-west region it has less measuring stations, so that those regions have comparatively smoothed grid values which were the rainfall corrected predictions from the long range atmospheric dispersal K-model. Also the variability among the data at grid points is less compare to variability on the data from measuring stations. For the EHBST model, on the region near München and Passau, the mean deposition is higher than the other two models, but simultaneously the posterior uncertainties were larger compare to the other two HBST models. The NWHBST model has the lowest uncertainty. This is because, in NWHBST model the covariance is learning from data and the daily data from measuring stations were available mainly from the southern Bavaria. Overall, the NWHBST model displayed the best posterior performances.

6.2.3 Validation of the Posterior Results Against Measurements

We ran the three different formulation of HBST models separately and it produced the posterior and predictive results for the parameters. Now, we need to validate these results. The posterior results from the EHBST, the SHBST and the NWHBST model has been validated against the measurements from 13 measuring sites. Table 6.6 gives the comparison of observed data (measurements) and the posterior means of the total near ground concentrations from the three HBST models for 30th April 1986 and 1st May 1986. All the posterior results

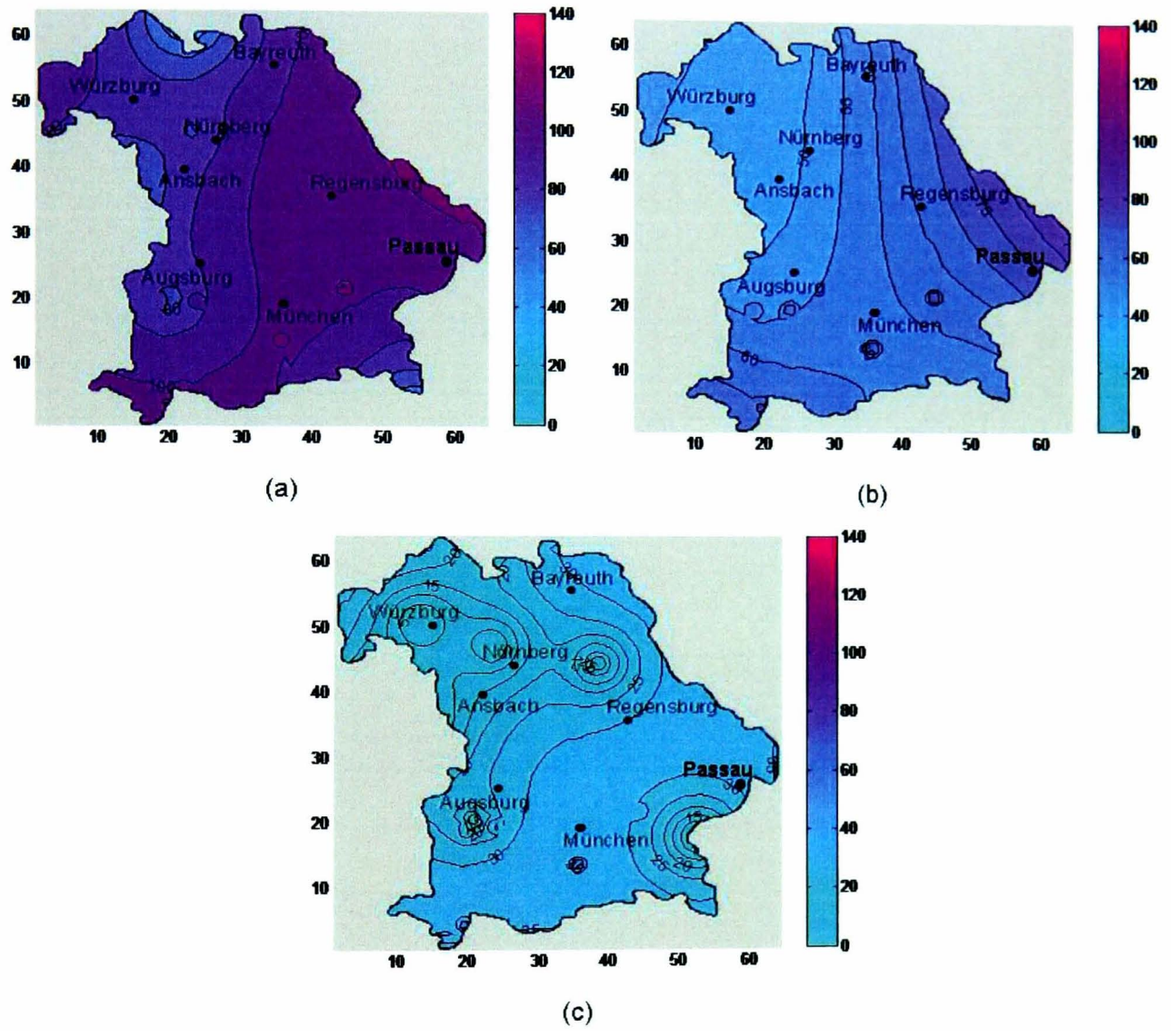


Figure 6.7: Posterior maps of mean radioactive deposition in Bavaria for (a) EHBST, (b) SHBST and (c) NWHBST models on 01 May 1986 (contours correspond to posterior means of the total radioactive deposition, based on the full data set)

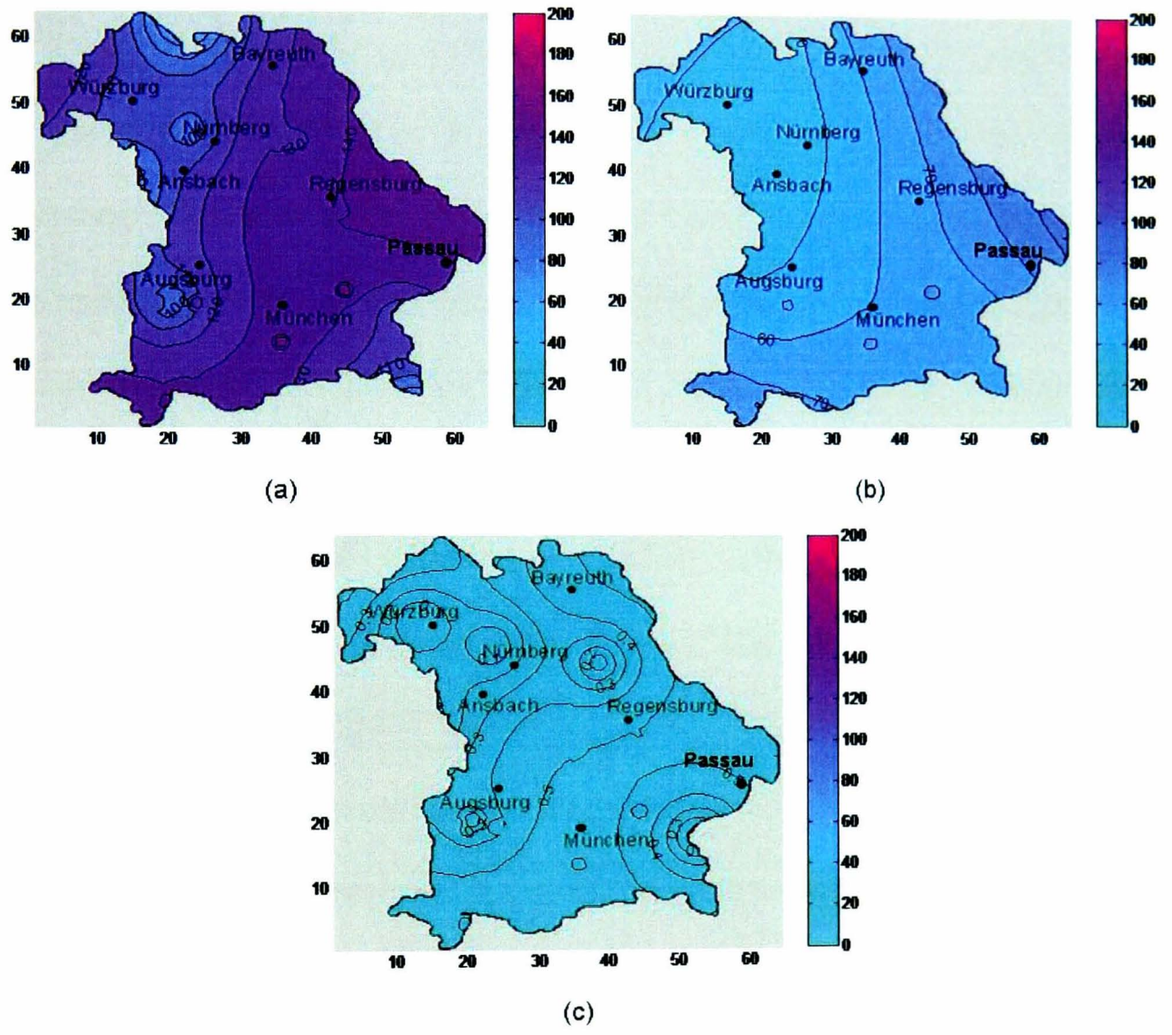


Figure 6.8: Posterior uncertainty maps in Bavaria for (a) EHBST, (b) SHBST and (c) NWHBST models on 01 May 1986 (contours correspond to posterior standard deviation of the total radioactive deposition, based on the full data set)

produced for the original variables of the total near ground concentrations. For all the HBST models, the probability distributions of the original variables are log-normal distribution as the transformed variable which were used in the HBST models are normally distributed.

Table 6.6 displays that, on 30th April and 1st May 1986 has the higher level of total near ground concentrations in the southern region of Bavaria (like, München, Ebersberg, Gundremmingen, Essenbach and Passau) and also small part of eastern region of Bavaria (Schwandorf). Table 6.6 also gives an idea of the performances of each of the HBST model based on comparison with the measurements. Both 30th April and 1st May, the posterior means of the total near ground concentrations for NWHBST model are very close to the measurement compare to the other two HBST models, the EHBST and SHBST.

To measure the model's 'best posterior fits' performance, we calculate the Spatial square-Root Mean Square Error at time t, called SRMSE(t), based on the 13 measuring sites for each of the HBST model separately in Table 6.5. The SRMSE(t) measures is based on only for fixed time point t. So, we calculate the SRMSE(t) for 30th April ($t = 5$) and 1st May ($t = 6$) 1986 separately. We can define SRMSE(t) in the following way.

$$SRMSE(t) = \sqrt{\frac{1}{m}\sum_{i=1}^{m} e_{it}^2} \tag{6.4}$$

where m is the total number of measuring sites and the e_{it} is the estimation error at measuring site i for the HBST model. The error e_{it} is the difference

between the measured and the posterior mean values of the total near ground concentrations from the HBST model.

Table 6.5: For 30th April 1986 ($t = 5$) and 1st May 1986 ($t = 6$), the SRMSE(t) values for the three HBST models

Day	EHBST	SHBST	NWHBST
30th April 1986	36.72	67.74	25.84
1st May 1986	42.52	53.49	24.94

From Table 6.5, we can see that for both $t = 5$ and $t = 6$, the SRMSE(t) values are less for the NWHBST model compare to the EHBST and the SHBST models. In other words, for the NWHBST model the posterior mean values for the total near ground concentrations are closer to the measurements. That is, NWHBST model gives the 'best posterior fits' to the measurements.

6.2.4 Predictive Distributions and the HBST Models Performances

The RAMSEs of all three models were computed to assess their 'short' (one-day-ahead) as well as their 'medium' term (three-days-ahead) predictive performances for the total near-ground radioactivity concentrations. That is, for each day t, the means of each model's predictive distribution of total concentrations at all underlying spatial points for days $t + 1$ and $t + 3$ were used to calculate the forecasting errors in the computation of the 'short' and 'medium' term RAMSEs respectively.

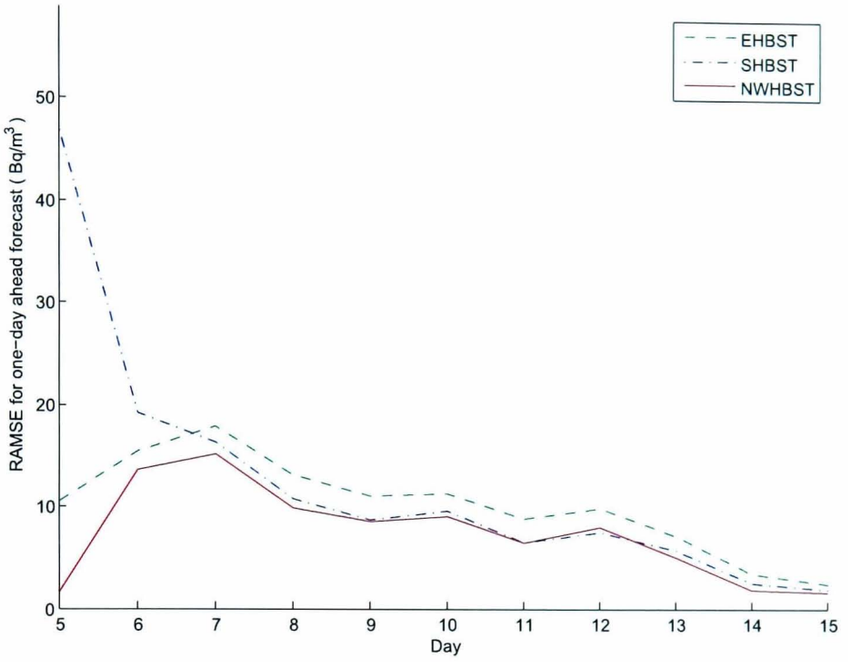
Figures 6.9(a) and (b) show the RAMSE plots for the one-day-ahead and the

Table 6.6: Observed and Posterior values of the original variables of total near-ground concentrations of the EHBST, the SHBST and the NWHBST models calculated on the 01 May 1986 at the 13 measuring stations (Egling (M_1), München (M_2), München-Ebersberg (M_3), Garching (M_4), Wasserburg (M_5), Gundremmingen (M_6), Essenbach (M_7), Niederaichbach (M_8), Passau (M_9), Schwandorf (M_{10}), Kahl (M_{11}), Grafenrheinfeld (M_{12}) and Hof (M_{13})).

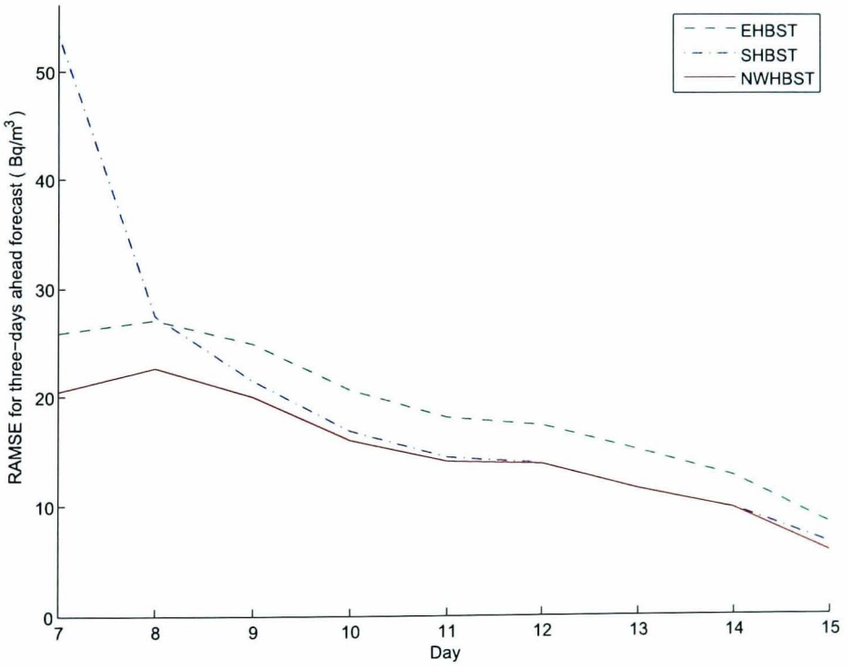
θ	30th April 1986				1st May 1986			
	Observed	Posterior Means			Observed	Posterior Means		
	Mean	EHBST	SHBST	NWHBST	Mean	EHBST	SHBST	NWHBST
θ_{M_1}	0.01	31.38	57.80	15.38	0.01	42.80	51.30	18.04
θ_{M_2}	79.58	67.08	126.30	33.56	64.13	103.90	85.70	39.54
θ_{M_3}	16.16	54.74	85.50	26.04	67.32	93.00	59.40	36.81
θ_{M_4}	0.71	19.34	31.30	12.30	5.56	32.90	25.60	17.05
θ_{M_5}	0.01	29.77	46.90	14.68	0.01	41.70	45.40	17.77
θ_{M_6}	69.95	82.52	146.90	34.54	76.58	109.40	153.00	30.97
θ_{M_7}	97.00	194.82	262.80	76.37	78.84	138.10	222.60	91.04
θ_{M_8}	0.01	28.94	72.80	15.30	22.00	78.90	50.10	35.31
θ_{M_9}	44.53	31.34	69.60	26.54	108.79	77.90	148.00	54.46
$\theta_{M_{10}}$	81.82	55.14	66.90	26.43	26.67	80.10	46.80	31.26
$\theta_{M_{11}}$	2.60	34.22	64.20	19.70	10.55	57.80	44.20	25.24
$\theta_{M_{12}}$	4.96	39.32	56.80	21.37	20.28	74.00	52.90	31.10
$\theta_{M_{13}}$	4.52	10.88	16.70	7.89	0.64	14.60	12.60	8.87

three-days-ahead forecasts respectively for each of the three models. Note that the NWHBST (solid line) had the best overall performance of the three models in both short and medium term forecasting. The Chernobyl contamination in Bavaria is measured from Day 4 (29th April 1986) to Day 15 (10th May 1986). So, the RAMSE for one-day ahead forecast starts from Day 5 (30th April 1986) as shown in Figure 6.9 (a) and the RAMSE for three-day ahead forecast starts from Day 7 (2nd May 1986) as shown in Figure 6.9 (b).

The concentration increased from practically natural levels to very high levels in the period from day 5 to day 7 for the short term, and on days 7 and 8 for the medium term forecasting, all the models' performances decreased with the RAMSE jumping to values in the range from 1.63 to 46.83 Bq/m^3 and from 20.50 to 53.19 Bq/m^3 for the short and medium respectively. However, in those periods the NWHBST produced the best relative performance of all models, with its RAMSE being on average 30.6% and 18.5% lower (for the short and the medium term periods respectively) than the second best model, the EHBST. The SHBST was the worst performing model during those periods, followed by the EHBST. In the period from day 8 to 15 and from day 9 to 15 (for the short and medium term periods respectively), all the models performances improved with their RAMSEs practically displaying a negative trend. Again, on those periods the NWHBST model slightly outperformed the remaining models. This time though the RAMSE from the NHWBST model was only 4.6% and 3.6% lower (for the short and medium term periods respectively) than the second best



(a)



(b)

Figure 6.9: The RAMSE for (a) one-day ahead forecasts, and (b) three-days ahead forecasts for the EHBST (dashed line), SHBST (dash-dotted line) and the NWHBST (solid line) models.

model, the SHBST. Overall, the NWHBST displayed the best performance on the RAMSE criteria showing average gains of 36.7% and 26.6% (for the short and medium term periods respectively) over the second best model, the EHBST.

The contour plots (a) to (c) in Figure 6.10 display the one-day ahead forecasts of total near-ground radioactivity air concentrations in Bavaria on 01 May 1986 by the three HBST models. We noticed that in general the one-day ahead predictions varied from $300Bq/m^3$ to $350Bq/m^3$ for the EHBST model, where as for the SHBST model it reached $600Bq/m^3$ to $700Bq/m^3$ at the region near München, Passau and Regensburg, i.e. the north-east and the south-east region of Bavaria. On the other hand, for the NWHBST model the one-day ahead prediction varied from $50Bq/m^3$ to $150Bq/m^3$ in the same region mentioned above. Comparing those forecasts with the measured values in some measuring stations located in that region shown in Figure 6.3, we notice that in general on 1st May 1986 the NWHBST model produced better results (i.e. close to the measured values) compare to the EHBST and SHBST models. As we have seen before, the RAMSE for the one-day-ahead forecast is larger on day 5 (1st May) for the SHBST model, which can also be seen from the contour plots for one-day ahead forecasts.

The contour plots (a) to (c) in Figure 6.11 show the corresponding predictive uncertainty maps for the total near ground concentrations in Bavaria on 01 May 1986 by the three HBST models. By incorporating this information, we can greatly improve the prediction provided by the one-day ahead prediction alone. Comparing those uncertainty maps, we notice that in general on 01 May 1986

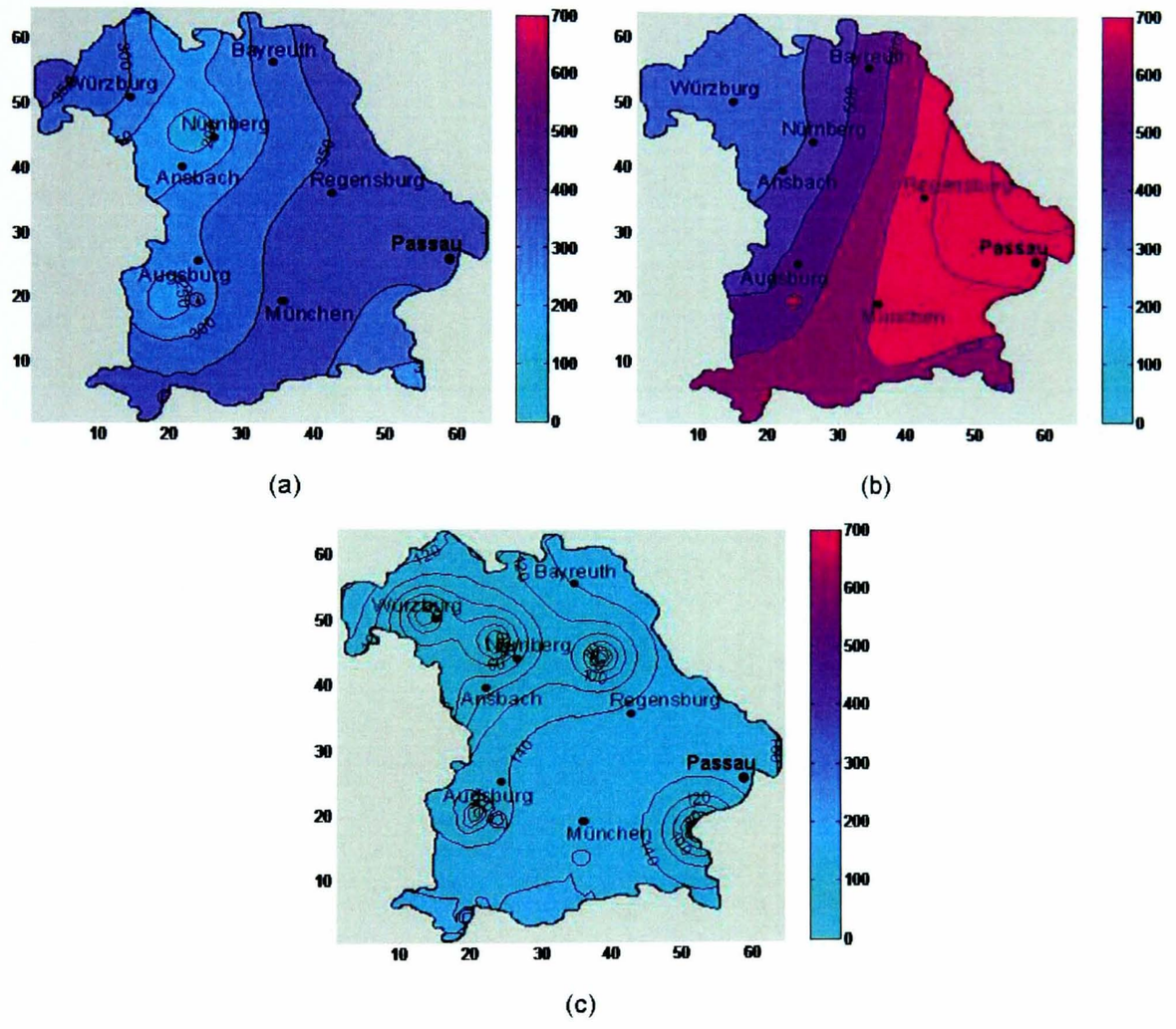


Figure 6.10: The one-day ahead predictive means in Bavaria for the (a) EHBST, (b) SHBST and the (c) NWHBST models on 01 May 1986 (contours correspond to predictive means of the total near ground radioactive deposition, based on the full data set)

the predictive uncertainties were larger for the EHBST model compare to other two models. The NWHBST model has the lowest uncertainty which can be seen from the lowest RAMSE for one-day ahead forecast in Figure 6.9 (a).

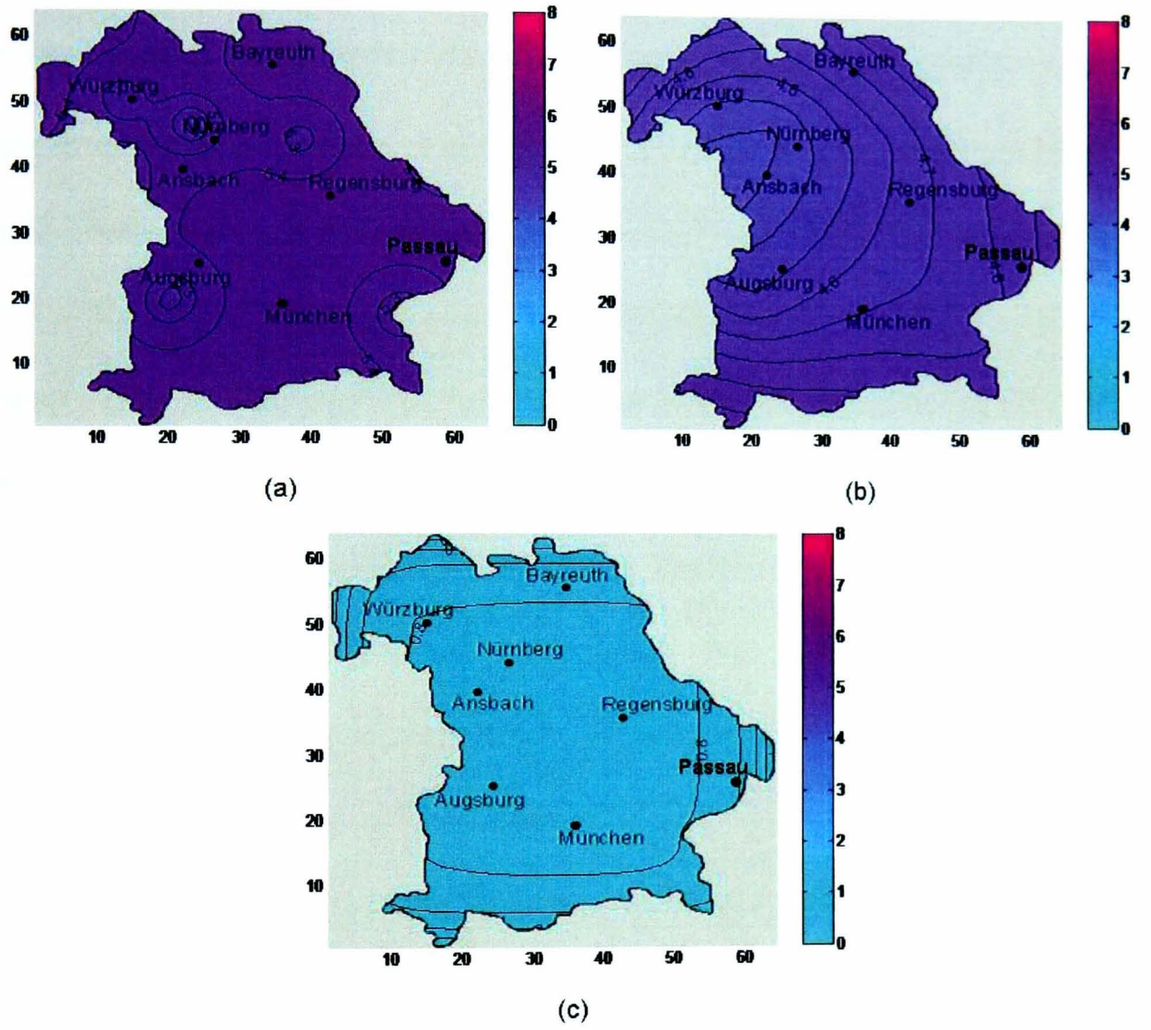


Figure 6.11: The one-day ahead predictive uncertainty maps in Bavaria for the (a) EHBST, (b) SHBST and the (c) NWHBST models on 01 May 1986 (contours correspond to predictive standard deviation of the total near ground radioactive deposition, based on the full data set)

6.2.5 Radionuclide Specific Predictive Performances of the HBST models for Decision Support for Nuclear Emergencies

The bar charts (a) to (f) in Figure 6.12 display for each day from the 29th of April (day 4) to the 4th of May 1986 (day 9), the one-step-ahead forecasts of near-ground radioactivity air concentrations of iodine (I131), caesium (Cs134 and Cs137) and others (Ru103, Ru106 and Te132) by the NWHBST model for the sites of the 13 measuring stations. Comparing those forecasts with the measured values shown in Figure 6.3, we notice that in general for each day the NWHBST model produced mixed results with under and over forecasting at some sites but also some pretty close predictions at others. For the 30 April and 01 May 1986, the predicted concentrations were a lot larger than on the 29 April and closer to the measured values at some sites, in particular for sites 6 and 7. This is because there were more accurate rainfall predictions near those sites. Such that, the rainfall measuring stations Augsburg and Ingolstadt are very close to the radioactivity measuring stations Gundremmingen (site 6) and Essenbach (site 7) respectively. Due to same reason of inaccurate rainfall information, the NWHBST model produced mixed results with under and over predictive values in some measuring sites.

Note that on day 5, the predicted level of caesium was larger than the measured at site 6 but not by a large amount, and at site 7, the predicted caesium level was smaller than the measured but again by a small amount. On day 6, the

predicted level of caesium varied in a pattern similar to day 5 at the sites 6 and 7. In general at site 6, the predicted caesium level was larger than the measured but by small amounts between days 5 and 9.

The predicted levels of iodine for sites 6 and 7 were very close to the measurements on days 5 and 6. At Munich (site 2), the predicted level of iodine was very close to the measurements between days 5 and 9. In general, on days 5 and 6, the predicted levels of iodine were very close to the measurements at most of the measuring sites except at 1, 5 and 8.

As the average natural level of concentration for caesium was $0.0009 Bq/m^3$ and the NWHBST models' predicted level of concentrations of caesium for site 6 (Gundremmingen) and site 7 (Essenbach) were $12.49 Bq/m^3$ and $3.82 Bq/m^3$ on 30th April and $15.84 Bq/m^3$ and $4.56 Bq/m^3$ on 1st May. So the level of concentrations of caesium for sites 6 and 7 were approximately 13878 and 4244 times larger on the 30th April, and 17600 and 5066 times larger on the 1st May than the natural levels of concentration.

On the other hand, as the average natural level of concentration for iodine was $0.0009 Bq/m^3$ and the NWHBST models' predicted level of concentrations of iodine for sites 6 and 7 were $38.66 Bq/m^3$ and $44.15 Bq/m^3$ on the 30th April and $30.21 Bq/m^3$ and $47.99 Bq/m^3$ on the 1st May. So, the predicted level of concentrations of iodine for sites 6 and 7 were approximately 42955 and 49055 times larger on 30th April and 33566 and 53322 times larger on the 1st May than the natural levels of concentration.

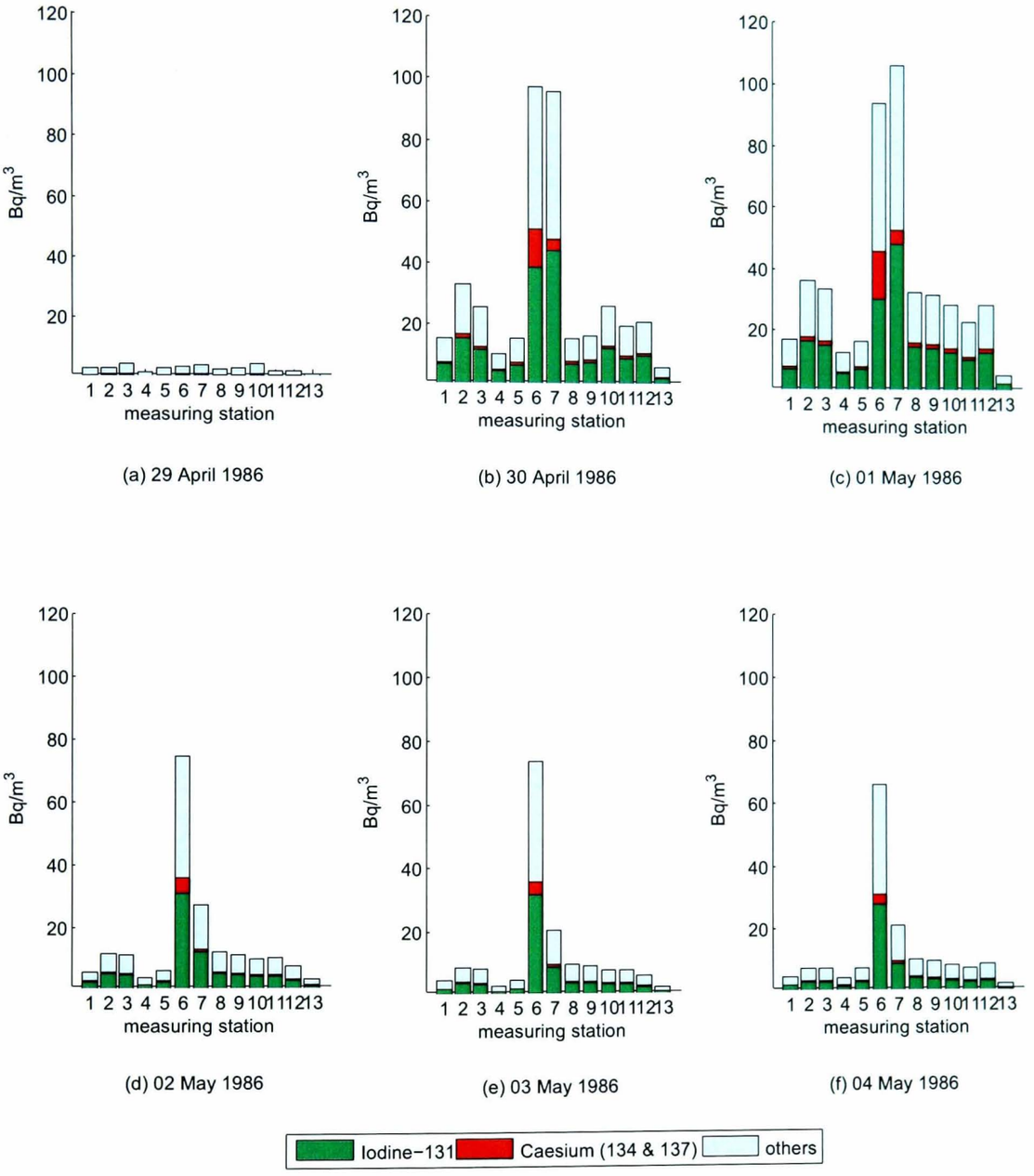


Figure 6.12: NWHBST model's predicted near-ground daily average air concentrations of iodine (I131), Caesium (Cs134 and Cs137) and other less hazardous isotopes (Ru103, Ru106 and Te132) from the 29 April ($t = 4$) to the 4 May 1986 ($t = 9$) at each of the 13 measuring stations in Bavaria: 1. Egling, 2. München, 3. München-Ebersberg, 4. Garching, 5. Wasserburg, 6. Gundremmingen, 7. Essenbach, 8. Niederaichbach, 9. Passau, 10. Schwandorf, 11. Kahl, 12. Grafenrheinfeld and 13. Hof.

From a decision maker's point of view, it may be important to predict the concentration levels at urban areas of larger populations. So we have produced one-day ahead predictions by the NWHBST model for the main cities in Bavaria. Table 6.7 shows the one-day ahead predicted daily average air concentrations (in Bq/m^3) and the corresponding observed (or, interpolated) values for the iodine and caesium isotopes calculated on the 30th April and 1st May 1986 at the five Bavarian cities: Augsburg, München, Ingolstadt, Nürnberg, Würzburg. To obtain the interpolated values at those cities, we adopted an interpolation rule similar to that used by the EHBST model in the previous section but with a smaller spatial decay of 0.0189 for the exponential form of the elements of the observational covariance matrix.

The large cities and the corresponding nearest measuring stations in parenthesis are: Augsburg (Gundremmingen), München (München), Ingolstadt (Essenbach), Nürnberg (Schwandorf), Würzburg (Grafenrheinfeld). These cities mainly cover the region of south-west (Augsburg), south (München), south-east (Ingolstadt), north-east (Nürnberg) and north-west (Würzburg) in Bavaria.

Note that, the predicted values of iodine and caesium by the NWHBST model fit well the observed (or interpolated values) shown in Table 6.7. The values with '$*$' marks lie outside the 95% predictive interval.

Table 6.7: The NWHBST model's one-day ahead predicted (pred) daily average air concentrations (Bq/m^3) and the corresponding observed/interpolated (obs/int) values (Bq/m^3) for the two groups of radionuclides (Iodine and Caesium) calculated on the 30th April and 1st May 1986 at some of the large cities (Augsburg, München, Ingolstadt, Nürnberg, Würzburg) in Bavaria.

Major Cities	30th April 1986				1st May 1986			
	Iodine		Caesium		Iodine		Caesium	
	obs/int	pred	obs/int	pred	obs/int	pred	obs/int	pred
Augsburg	31.03	31.56	3.61	10.19*	28.94	24.67	6.31	12.93*
München	3.04	11.55	1.28	1.00	15.36	12.42	6.67	1.18*
Ingolstadt	5.95	7.80	1.38	0.68	5.24	8.49	1.29	0.81
Nürnberg	2.12	1.04	0.79	0.09	0.86	1.12	0.24	0.11
Würzburg	0.26	3.15	0.25	0.27	2.62	4.18	0.36	0.39

6.3 Model Sensitivity

This section describes the HBST model sensitivity by the source of variation in the input parameters of the three HBST models. In the following subsections I will discuss only the sensitivity of the three HBST models to the K-model predictions and the sensitivity of the spatial decay in the observational covariance matrix $V_{Z_{Dt}}$ in equation (5.14) of the EHBST model. Note that, I have restricted the sensitivity of the spatial decay in the observational covariance matrix only for EHBST model, and not for the SHBST model, as there is no parameter (like the decay parameter) in the SHBST model that can vary.

6.3.1 Sensitivity of K-Model

This subsection describes the variation in the output of the HBST model to the sources of variation in the input parameters of the K-model. The K-model is

simple Eulerian diffusion model for long-range atmospheric transport of activity. The parameters of the K-model equation in (6.1), the total activity removal rate (Γ), the deposition velocity for each isotope j (ν_j) and the large scale eddy diffusivity coefficient (K) are to be fitted by the experiment.

In our application, we adopted the values of $K = 10^6 m^2 s^{-1}$ and $\Gamma = (10 days)^{-1}$ or $\Gamma = (864000)^{-1} s^{-1}$ adopted by Lauritzen and Mikkelsen (1999) for the Chernobyl accident. The prediction from K-model is highly negatively correlated with the value of K. That is, as we increase the value of K, the near ground predicted concentrations from K-model decrease rapidly at the grid points.

Also in our application we used the average wind speed projecting towards Bavaria at Chernobyl. The accurate wind speed and directions will produce more accurate K-model predictions at the grid points which will give the more accurate predicted near ground air concentrations for the HBST models.

The K-models' predicted concentrations at the grid points are positively correlated with the deposition velocity of each isotope. In our application we assumed that the deposition velocity is constant for all the radionuclides. The different values of the deposition velocity for different isotope will give more accurate nuclide specific predictions for the K-model.

In general, as the value of the K-model predictions increases at the grid point, which gives a large predictive interval for the HBST model's predicted near ground concentration at the grid points as well as in the measuring stations.

6.3.2 Sensitivity of Spatial Decay for Spatial Covariances

In the EHBST model, the elements of the observational covariance matrix in equation (5.14) are *known* and *fixed* by an exponential form. So, to check the EHBST model sensitivity, we used different values of the spatial decay $\eta_{c_{ijt}}$ in equation (5.14). In the application in Chapter 5, we used the same spatial decay independently of spatial locations and time. Also, this section shows the sensitivity analysis for the total near ground radioactivity deposition only, i.e. when $c = Z$ in equation (5.14). For simplicity, from now $\eta_{c_{ijt}}$ will be denoted as η_Z for $c = Z$.

The predictive distribution is quite sensitive to slight changes of the spatial decay η_Z. Figure 6.13 shows the contour plots of one-step ahead prediction and the corresponding uncertainty maps for $\eta_Z = 0.005$, 0.01, 0.0189 and 0.05 respectively.

Figure 6.14 shows the RAMSE plots for the one-day ahead forecasts respectively for each value of decays described above. For the above mentioned values of η_Z, we note that in Figure 6.14, the RAMSE values were increased as the values of decay increased. In our application for the EHBST model, we estimated the decay parameter from the data as $\eta_Z = 0.0189$. We also note that, as the value of spatial decay increased the one-day ahead predicted near ground radioactivity concentrations and the corresponding predictive uncertainties were also increased which can be seen from the Figure 6.13.

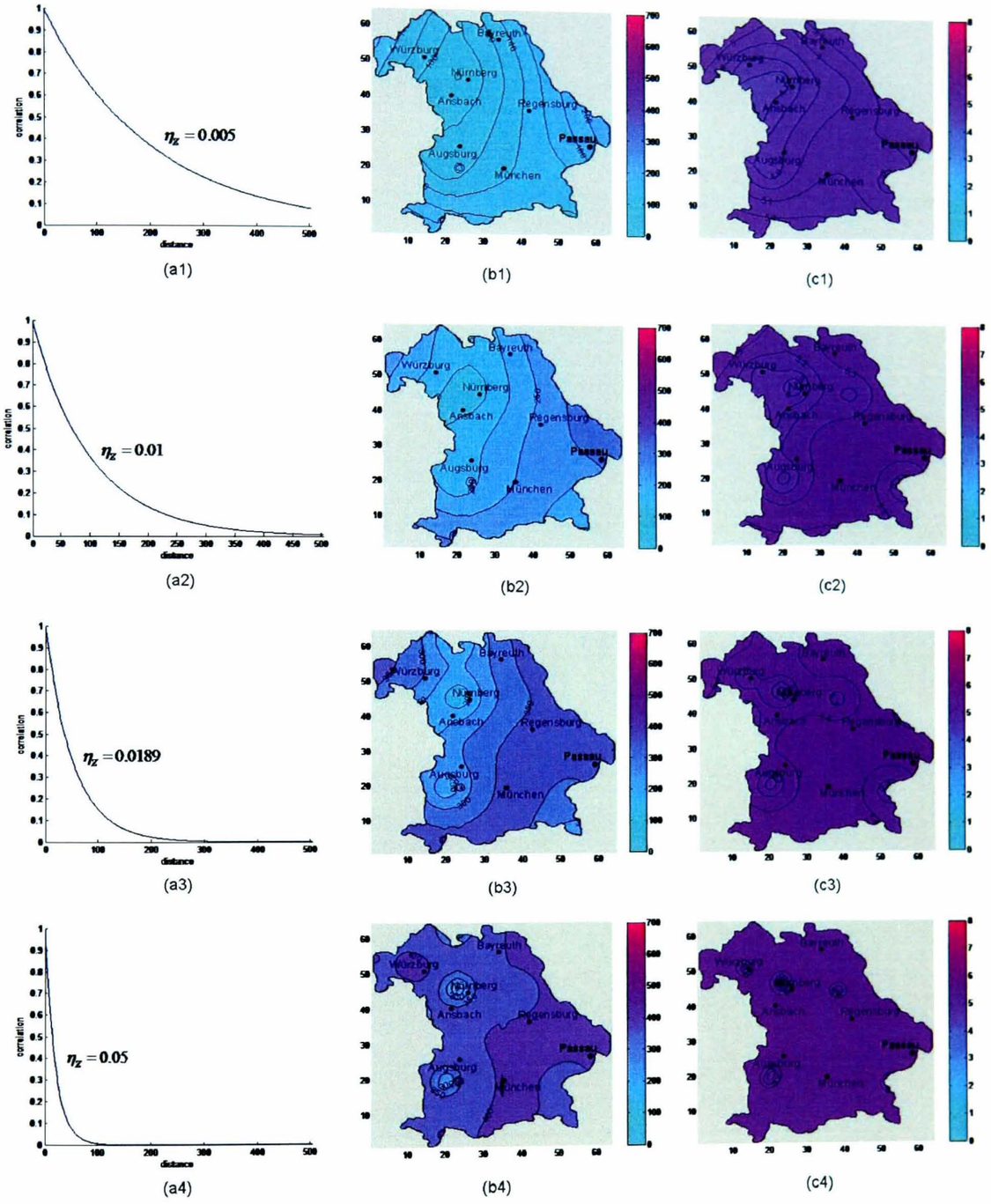


Figure 6.13: Contour plots of the one-day ahead predictions and the corresponding uncertainty maps for different spatial decays of the EHBST model on 01 May 1986. The plots of different decays (viz. $\eta_Z = 0.005, 0.01, 0.0189, 0.05$) are shown in (a1), (a2), (a3), and (a4) respectively. The one-day ahead forecast and the uncertainty maps are shown in (b1), (b2), (b3), (b4) and (c1), (c2), (c3), (c4)) respectively, corresponding to each value of η_Z.

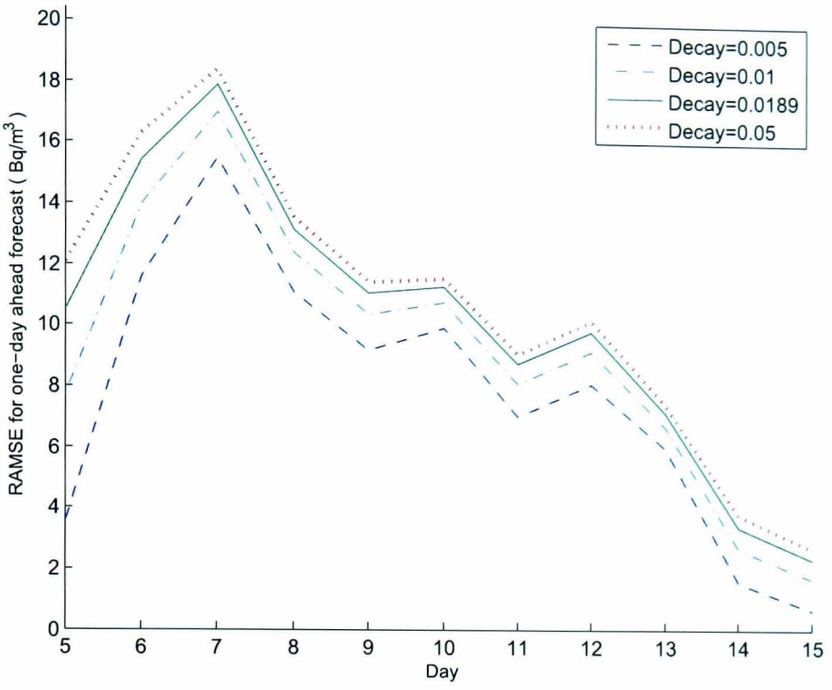


Figure 6.14: The RAMSE for one-day ahead forecast for different decays ($\eta_Z = 0.005, 0.01, 0.0189, 0.05$) of the EHBST model.

Figure 6.14 shows that, for the EHBST model, the smaller value of spatial decay gives the the smaller RAMSE, or in other words, better predictive results. But in our application we estimated the decay from the data as 0.0189 . This is because the smaller value of decay produced spatially smoothed predictions from the K-model, in their turn, would degrade the quality of the EHBST model's predictions.

The contour plots (b1) to (b4) in Figure 6.13 show that the smaller decay value ($\eta_Z = 0.005$) produced the smoothed predictions from the EHBST model in Bavaria, and in general, it varied between $100Bq/m^3$ and $120Bq/m^3$, and slightly larger near Passau (approximately $200Bq/m^3$). Then at $\eta_Z = 0.01$, the

variation in the near ground air concentrations increased between $100Bq/m^3$ and $250Bq/m^3$ in the whole Bavaria, except near Passau (approximately $300Bq/m^3$). The variation reached between $100Bq/m^3$ and $350Bq/m^3$ as we used the value of $\eta_Z = 0.0189$, which produced better results (i.e. close to the variation present in the measured values) compare to other values of η_Z.

6.4 Cross-Validation of HBST Models

In subsections 6.2.4 and 6.2.5, the predictive performances of the HBST model have already been described. This section describes the predictive performances of the HBST models in another way, called, cross-validation technique. As the NWHBST model performed better among all the other HBST models, so in this section we concentrated only on the NWHBST model. That is the cross-validation technique was applied to asses the NWHBST model's predictive performance.

To assess the model's predictive performance, we split the data into two parts: a training sample which consists 10 measuring stations and the other part is validation set, which consists 3 measuring stations. The measuring stations under training sample are Egling, München-Ebersberg, Garching, Wasserburg, Niederaichbach, Passau, Schwandorf, Kahl, Grafenrheinfeld and Hof. The measuring stations in the validation set are München, Gundremmingen and Essenbach. Both the data set have to be representative samples of the data that the NWHBST model would be applied to.

The cross-validation technique applies to asses the model's performance for an independent data set. This is mainly use to check the model's predictive performance. In other words, to estimate the accuracy of the model in predictive sense. In Cross-validation technique one question arises regarding the splitting the samples. Picard and Cook (1984) suggested a well-motivated procedure to split the data at random. According to Picard and Cook (1984), if the original data can be viewed as a sample randomly drawn from a population of interest, a randomly selected validation set should mimic a sample of future observations since such a split is easy to implement. But in our case the selected samples in the validation set were not random. They were selected in such a way so that the samples in the validation set were close to the samples in the training set as we shall see in the following subsection.

6.4.1 Training Sample and Validation Set are Close to Each Other

We have chosen each sample in the validation set close to the sample(s) in the training set. E.g. from Figure 6.1 we can see that, München is close to Egling and München-Ebersberg, Gundremmingen is close to Wasserburg, and Essenbach is close to Niederaichbach. The reason behind these choices were to minimise the predictive errors as we have only 13 sparse fixed measuring stations in Bavaria. We computed the predictive distribution for the 3 measuring stations which belongs to validation set. Table 6.8 shows the one-day ahead predictive total near ground air concentrations (in Bq/m^3) by the NWHBST model and the corre-

sponding observed values of the 3 measuring stations in the validation set on 1st May 1986. The NWHBST model produced an over forecasting at those measuring stations in the validation set. This is common for cross-validation technique as we shall see.

Table 6.8: The observed and the one-day ahead predicted total near ground air concentrations by the NWHBST model at the 3 measuring stations (viz. München, Gundremmingen and Essenbach) in the validation set on 1st May 1986

Measuring	1st May 1986	
Station	observed	predicted
München	64.13	139.36
Gundremmingen	76.58	193.16
Essenbach	78.84	199.48

The calculated square-Root of the Mean Square Error (RMSE) values for each measuring station in the training sample as well as in the validation set are shown in the Figure 6.15. The RMSE is calculated by

$$RMSE = \sqrt{\frac{1}{\tau}\sum_{t=1}^{\tau}\left(\mathbf{Y}_{it} - \widehat{\mathbf{Y}}_{it}\right)^2} \tag{6.5}$$

where $\mathbf{Y}_{it}$ and $\widehat{\mathbf{Y}}_{it}$ are the observed and the one-step ahead predicted total near ground air concentrations respectively for the ith sample (either from training set or validation set). τ is the total number of release days. Note that, the RMSE measures the mean difference between actual and predicted values.

In Figure 6.15, the 'red circles' denote the RMSE values for each sample in the training set and the validation set. Also, the first 10 samples (measuring

stations) were from the training sample and the last 3 samples (after the blue vertical line) were from the validation set. The vertical line (blue) represents the end of the training sample.

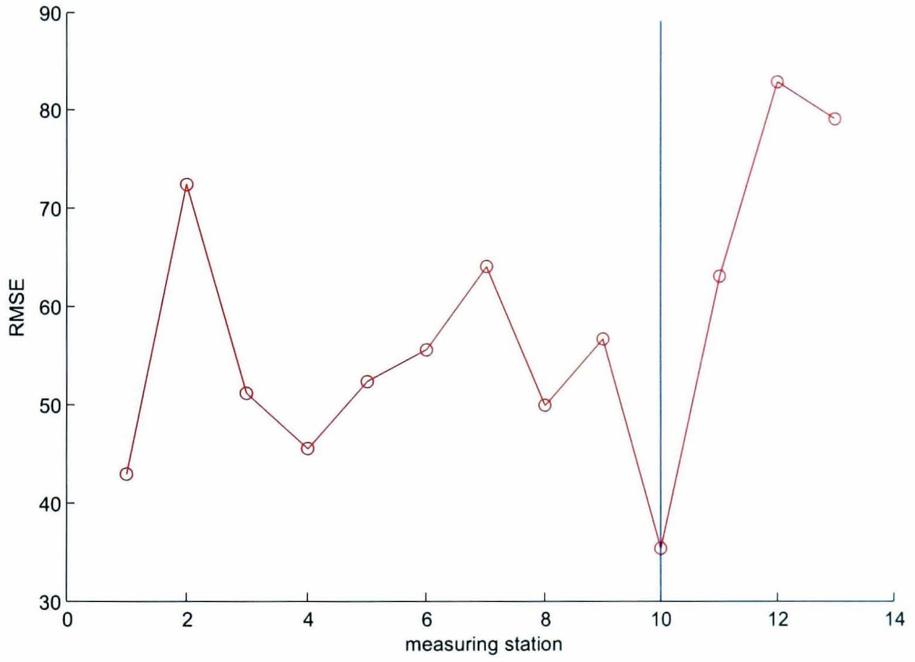


Figure 6.15: The plots of RMSE values for training sample & validation set: the vertical line (blue) represents the end of the training sample.

Note that the RMSE values for the 3 measuring stations (viz. München, Gundremmingen and Essenbach) in the validation set were slightly larger than the RMSE values in the training sample. The average RMSE of validation set was 44% larger than the average RMSE of the training sample. This is normal to get slightly larger RMSE for the validation set, as in the updating algorithm in Appendix B we used the measured values only for the training samples, not for the validation set. Also the other reason for the over forecasting would be the smaller size of the training sample.

In general if we take an independent sample of validation data from the same population as the training sample data, it will turn out that the model does not fit the validation data as well as it fits the training data. This is called overfitting, and is particularly likely to happen when the size of the training sample is small, or when the number of parameters in the model is large. No matter how the performance of the model is measured, we always need to measure on the validation set, not on the training set. This is because the performance on the training sample only tells us that the model learns what it is supposed to learn. It is not good indicator of performance on validation set.

Finally we can conclude that the overall predictive performance of the NWHBST model is good though it produced slightly larger predictions than the measured values for the validation set but not by a large amount. Due to the smaller size of the training sample the accuracy was not good in our case, but the larger size of the training sample would give better predictions for the NWHBST model.

The next chapter describes a more general conclusion on the proposed HBST models and the direction of the possible ways to improve the HBST models.

Chapter 7

Conclusion and Direction for Further Research

In Chapters 4 and 5 we described the general framework of the proposed HBST models and an application of the proposed models to the radioactivity deposition. Then in Chapter 6 we analysed the results from the HBST models for the radioactivity deposition. This chapter summarises the main results of this thesis and propose directions for further research in this area.

7.1 Conclusion

This section describes the major contribution of this study. The HBST models proposed in this thesis provide a flexible framework for modelling nuclide specific near-ground radioactive concentrations in time and space. Those models are valid for any data sets that are discrete in time and continuous in space, including data observed at irregular grids or at locations that change over time. They can be adapted for use in different countries with different measuring resources. They use predictions from a long range atmospheric dispersal model as data to

produce improved estimation of concentration levels which account for weather information.

The accuracy of the HBST models depend crucially on the quality of inputs from the long range atmospheric dispersal model as this in principle accounts for atmospheric effects. Processes of the type being modelled here are meteorology dependent by nature. In our case, too smoothed predictions from the atmospheric model can lead to HBST models predictions with smaller associated uncertainties than expected being obtained. The K-model did a decent job here but other more sophisticated models such as RIMPUFF dispersion model by Mikkelsen *et al.* (1984), which explicitly accounts for atmospheric forecasts and measurements including temperature, precipitation and wind fields, are expected to perform better and thus improve the HBST models' predictive qualities.

The model formulation presented in this thesis, addresses the issue of compatibility between outputs of modules in a chain system that models all the phases of the whole accident and the inputs required by other modules further down in the chain. In particular, it makes it possible to coherently transfer information (updated by measurements and/or other information) between a long range atmospheric dispersal model and a food contamination model.

Finally, we can conclude that the proposed NWHBST model among all the HBST models that were considered performed better in the posterior and the predictive senses. Still the predictions from the NWHBST model could have been improved by incorporating more accurate rainfall information over the ge-

ographic region of interest. More accurate rainfall information would produce more accurate predictions by the K-model at the grid points which, in their turn, would improve the quality of the NWHBST's predictions. In this application we had only 13 sparse measuring sites in the Bavarian region to cover an area of approximately $70,549 Km^2$ as shown in Figure 6.1. Obviously, more measurements would contribute to better data assimilation and model performance.

One criticism of the K-model predictions is, the use of this release profile corresponds to using hindsight information (e.g. amount of the exact release materials). This certainly favors the K-model performance and consequently the HBST models applied in this thesis. In reality, no one will know the release profile for sure when an accident happens and the K-model to be run will need some estimation of that which is likely to be worse than a posterior estimation based on measurements of deposited radioactivities. Thus, in a real future accident the HBST models would certainly produce worse performances than the ones portrayed in this thesis. Perhaps a solution to this would be to choose prior parameters for the HBST models such as to weight down the K-models inputs when crude estimates are used.

7.2 Future Research

This thesis has not exhausted all the possible ways to improve the HBST models. This section describes the direction of further extension of the HBST models.

An extension of the proposed HBST model to include explanatory variables is

technically straightforward. It would only require the inclusion of a corresponding regression matrix in the set of the model's equations and consequently on the set of updating Kalman filter equations. Similarly, a generalization to distribution-free models (associated to processes that can be characterized by the first two moments only) on the line of Goldstein (1976) Bayes linear method would be relatively easy to implement.

Certainly a future step in our developments will consist of making use of a fully computational MCMC approach. A possible advantage of that would be a better quality of the posterior densities estimation achieved by the inclusion of the process non-linearities in the methodology. However, in this non-linear setting complications associated with unidentifiability and convergence are likely to be significant. Other possible extensions of our model include the use of anisotropic covariance functions as proposed by Sampson and Guttorp (1992) to consider directional spatial dependencies.

Off-grid unmonitored sites of interest (such as large cities, rivers, agricultural regions, etc) could also be included in the model by augmentation of vectors and matrices only. This would allow, via model interpolation, the estimation of more accurate contamination levels at those sites.

The proposed models can be used for decision support purposes with clear decision making issues potentially addressed (such as evacuation, relocation, food ban, etc.) although the HBST model may not necessarily be known to a decision maker.

Appendix A

A space-time extension of the Kalman filter algorithm for the HBST model

A.1 The initial information:

The distribution of $\boldsymbol{\Theta}_0 = (\boldsymbol{\Theta}_{D0}, \boldsymbol{\Theta}_{M0})'$ at time $t = 0$ is assumed to be a multivariate normal distribution with parameters $\mathbf{a}_0$ and R_0, that is

$$\boldsymbol{\Theta}_0 \sim N[\mathbf{a}_0, R_0]$$

where $\mathbf{a}_0 = (\mathbf{m}_0, G\mathbf{m}_0)'$ and $R_0 = \begin{pmatrix} C_0 & GC_0 \\ C_0G' & GC_0G' + \Sigma_0 \end{pmatrix}$

A.2 The prior distribution of $\boldsymbol{\Theta}_t$:

The prior distribution of $\boldsymbol{\Theta}_t = (\boldsymbol{\Theta}_{Dt}, \boldsymbol{\Theta}_{Mt})'$ at time t $(t = 1, 2, ...)$, conditional on the information $\mathbf{I}_{t-1}$ available at time $t-1$, is given by

$$(\boldsymbol{\Theta}_t | \mathbf{I}_{t-1}) \sim N[\mathbf{a}_t, R_t]$$

where $\mathbf{a}_t = (H_t\mathbf{m}_{t-1}, GH_t\mathbf{m}_{t-1})'$

and

$$R_t = \begin{pmatrix} H_t C_{t-1} H_t' + W_t & G H_t C_{t-1} H_t' + G W_t \\ H_t C_{t-1} H_t' G' + W_t G' & G H_t C_{t-1} H_t' G' + G W_t G' + \Sigma_t \end{pmatrix}$$

A.3 The one-step predictive distribution of $\mathbf{Y}_t$:

The one-step ahead predictive distribution of $\mathbf{Y}_t = (\mathbf{Y}_{Dt}, \mathbf{Y}_{Mt})'$ conditional on $\mathbf{I}_{t-1}$, is a multivariate normal distribution with mean vector $\mathbf{f}_t$ and covariance matrix Q_t, that is

$$(\mathbf{Y}_t | \mathbf{I}_{t-1}) \sim N[\mathbf{f}_t, Q_t]$$

where $\mathbf{f}_t = \mathbf{a}_t$ and $Q_t = R_t + V_t$

A.4 The posterior distribution of Θ_t:

The distribution of $\mathbf{\Theta}_t = (\mathbf{\Theta}_{Dt}, \mathbf{\Theta}_{Mt})'$ conditional on $\mathbf{I}_t = \{\mathbf{I}_{t-1}, \mathbf{Y}_t\}$ is a multivariate normal with mean vector $\mathbf{a}_t^*$ and covariance matrix R_t^*, that is

$$(\mathbf{\Theta}_t | \mathbf{I}_t) \sim N[\mathbf{a}_t^*, R_t^*]$$

where $\mathbf{a}_t^* = \mathbf{a}_t + A_t \mathbf{e}_t$ and $R_t^* = R_t - A_t Q_t A_t'$,

with $A_t = R_t Q_t^{-1}$, and $\mathbf{e}_t = \begin{pmatrix} \mathbf{Y}_{Dt} - H_t \mathbf{m}_{t-1} \\ \mathbf{Y}_{Mt} - G H_t \mathbf{m}_{t-1} \end{pmatrix}$ being the one-step-ahead forecasting error vector.

A.5 The k-step predictive distribution for $\mathbf{Y}_{t+k}$:

The k-steps ahead predictive distribution for $\mathbf{Y}_{t+k}$ given $\mathbf{I}_t$ is a multivariate normal with mean $\mathbf{f}_t(k)$ and covariance matrix $Q_t(k)$ where $k = 1, 2, ...$ That is.

$$(\mathbf{Y}_{t+k}|\mathbf{I}_t) \sim N[\mathbf{f}_t(k), Q_t(k)]$$

where $\mathbf{f}_t(k) = \mathbf{a}_t(k)$ and $Q_t(k) = R_t(k) + V_{t+k}$,

with $\mathbf{a}_t(k) = H_{t+k}\mathbf{a}_t(k-1)$ and $R_t(k) = H_{t+k}R_t(k-1)H'_{t+k} + W_{t+k}$,

where $\mathbf{a}_t(0) = \mathbf{a}_t^*$ and $R_t(0) = R_t^*$,

with $\mathbf{a}_t^*$ and R_t^* are defined in A.4.

Appendix B

The prior-to-posterior updating algorithm for the Normal inverse-Wishart HBST model

Suppose, at time t, the state parameter vector $\boldsymbol{\Theta}_t = (\boldsymbol{\Theta}_{Dt}, \boldsymbol{\Theta}_{Mt})'$ and the observational covariance matrix V_t are *unknown*.

B.1 The initial information:

Let, the distribution of $\boldsymbol{\Theta}_0 = (\boldsymbol{\Theta}_{D0}, \boldsymbol{\Theta}_{M0})'$ conditional on V_0 and on the initial information $\mathbf{I}_0$ be a multivariate normal with *fixed* and *known* hyperparameters $\mathbf{a}_0 = (\mathbf{m}_0, G\mathbf{m}_0)'$ where $\mathbf{m}_0 = (m_{D10}, \ldots, m_{Dn0}, \ldots, m_{D(k+2)n0})'$ are initial means at grid points for each component Z, P and Q, and $R_0 = \begin{pmatrix} C_0 & GC_0 \\ C_0G' & GC_0G' + \Sigma_0 \end{pmatrix}$, where C_0 of dimension $((k+2)n \times (k+2)n)$ is the initial spatial covariance matrix (similar to V_{Dt} defined in Subsection 5.2.3) associated with the components Z, P and Q for all the grid points.

$$(\boldsymbol{\Theta}_0 | \, V_0, \mathbf{I}_0) \sim N\left[\mathbf{a}_0, R_0\right]$$

and the distribution of V_0 conditional on the initial information $\mathbf{I}_0$ is an inverse-Wishart with the *fixed* parameters $d_0 \geq (k+2)(m+n)$ and $\Psi_0 > 0$,

$$V_0|\mathbf{I}_0 \sim IW\left[\Psi_0, d_0\right]$$

with $V_0 = \Psi_0/d_0$.

B.2 The prior distribution of $\mathbf{\Theta}_t$:

The prior distribution of $\mathbf{\Theta}_t = (\mathbf{\Theta}_{Dt}, \mathbf{\Theta}_{Mt})'$ at time t $(t \geq 1)$, conditional on V_0 and the information $\mathbf{I}_{t-1}$ is given by

$$\mathbf{\Theta}_t|\, V_0, \mathbf{I}_{t-1} \sim N\left[\mathbf{a}_t, S_0 R_t S_0\right]$$

and the distribution of V_t conditional on $\mathbf{I}_{t-1}$ is an inverse-Wishart with parameters $d_{t-1} \geq (k+2)(m+n)$ and $\Psi_{t-1} > 0$,

$$V_t|\mathbf{I}_{t-1} \sim IW\left[\Psi_{t-1}, d_{t-1}\right]$$

where $\mathbf{a}_t = (H_t\mathbf{m}_{t-1},\, GH_t\mathbf{m}_{t-1})'$

and

$$R_t = \begin{pmatrix} H_t C_{t-1} H_t' + W_t & GH_t C_{t-1} H_t' + GW_t \\ H_t C_{t-1} H_t' G' + W_t G' & GH_t C_{t-1} H_t' G' + GW_t G' + \Sigma_t \end{pmatrix}$$

B.3 The predictive distribution of $\mathbf{Y}_t$:

The one-step ahead predictive distribution of $\mathbf{Y}_t = (\mathbf{Y}_{Dt}, \mathbf{Y}_{Mt})'$ made at time $t-1$ is a multivariate Student-t distribution with mean vector $\mathbf{f}_t$ and covariance matrix $S_{t-1}Q_tS_{t-1}$, that is

$$\mathbf{Y}_t|\mathbf{I}_{t-1} \sim T\left[\mathbf{f}_t, S_{t-1}Q_tS_{t-1}, d_{t-1}\right]$$

where $\mathbf{f}_t = \mathbf{a}_t$. $S_{t-1}^2 = \Psi_{t-1}/d_{t-1}$ and $Q_t = I + R_t$, with an identity matrix I.

B.4 The posterior distribution of Θ_t:

The distribution of $\mathbf{\Theta}_t = (\mathbf{\Theta}_{Dt}, \mathbf{\Theta}_{Mt})'$ conditional on V_t and on $\mathbf{I}_t$ is given by

$$(\mathbf{\Theta}_t|\, V_t, \mathbf{I}_t) \sim N\left[\mathbf{a}_t^*, S_tR_t^*S_t\right]$$

and the distribution of V_t conditional on $\mathbf{I}_t$ is an inverse-Wishart with parameters d_t and Ψ_t,

$$V_t|\mathbf{I}_t \sim IW\left[\Psi_t, d_t\right]$$

where $\mathbf{a}_t^* = \mathbf{a}_t + A_t\mathbf{e}_t$, $R_t^* = S_t(R_t - B_tQ_tB_t')S_t$,

$S_t^2 = \Psi_t/d_t$, $d_t = d_{t-1} + 1$,

and $\Psi_t = \Psi_{t-1} + S_{t-1}Q_t^{*-1}\mathbf{e}_t\mathbf{e}_t'Q_t^{*-1}S_{t-1}$,

with $Q_t = Q_t^{*2}$, $B_t = R_tQ_t^{-1}$ and $A_t = S_tB_tS_t^{-1}$,

and $\mathbf{e}_t = \begin{pmatrix} \mathbf{Y}_{Dt} - H_t\mathbf{m}_{t-1} \\ \mathbf{Y}_{Mt} - GH_t\mathbf{m}_{t-1} \end{pmatrix}$ being the one-step-ahead forecasting error vector.

B.5 The k-steps predictive distribution for $\mathbf{Y}_{t+k}$:

The k-steps ahead predictive distribution for $\mathbf{Y}_{t+k}$ given $\mathbf{I}_t$ is a multivariate Student-t distribution with mean $\mathbf{f}_t(k)$ and covariance matrix $S_{t-1}(k)\,Q_t(k)S_{t-1}(k)$ where $k = 1, 2, ...$ That is,

$$(\mathbf{Y}_{t+k}|\mathbf{I}_t) \sim T\left[\mathbf{f}_t(k), S_{t-1}(k)\,Q_t(k)S_{t-1}(k), d_{t-1}(k)\right]$$

where $\mathbf{f}_t(k) = \mathbf{a}_t(k)$ and $Q_t(k) = R_t(k) + I$,

with $\mathbf{a}_t(k) = H_{t+k}\mathbf{a}_t(k-1)$ and $R_t(k) = H_{t+k}R_t(k-1)H'_{t+k} + W_{t+k}$,

where $\mathbf{a}_t(0) = \mathbf{a}_t^*$ and $R_t(0) = R_t^*$,

and $S_{t-1}(0) = S_{t-1}$ and $d_{t-1}(0) = d_{t-1}$,

with $\mathbf{a}_t^*$, R_t^* are defined in B.4,

and S_{t-1} and d_{t-1} are defined in B.3.

Appendix C

Factorisation Method for Positive Definite Matrices

The two basic factorisation methods for positive definite (p.d.) matrices are considered. First, the Cholesky decomposition method which gives the square roots of a symmetric matrix. Second, the Jacob method which provide full eigen-structure, and consequently square roots also. Following are the basic definitions and properties of these two decomposition methods.

C.1 The Cholesky Decomposition and the Matrix Square Roots Algorithm

Definition C.1. If the positive definite matrix V can be written as $V = SS'$, with S a square matrix, we say that S is a square root of V. This definition, although frequently used, is not universal. Some authors allow S to be rectangular, while others restrict S to be symmetric.

Properties :

(i) Unicity : Square Roots matrices, when they exist, are non-unique. If S is

one square root of V, and P is orthogonal ($PP' = I$), then SP is also a square root of V.

(ii) Existence : Every p.d. matrix has a square root. this can be verified by the elementary process of completing squares in the corresponding quadratic forms.

(iii) Triangular Factorisation : If $S1$ and $S2$ are two lower triangular factorisations of V, then $S1 = S2$ diag$(\pm 1, \cdots \pm 1)$, i.e. $S(j,j) = \pm[V(j,j)]^{1/2}$. From now on, we will consider the unique square root corresponding to positive diagonal elements.

C.1.1 Lower Triangular Cholesky Decomposition

If V is p.d., it has a lower triangular factorisation $V = SS'$, and the one with positive diagonal elements is given by the following algorithm : For $j = 1, \ldots d-1$ (d = dimension of V), recursively cycle through the following ordered set of equations:

$$S(j,j) = V(j,j)^{1/2}$$

$$S(k,j) = V(k,j)/S(j,j), \quad k = j+1, \ldots d$$

$$V(i,k) \leftarrow V(i,k) - S(i,j)S(k,j),$$

with $k = j+1, \ldots d$ and $i = k, \ldots d$, and then $S(d,d) = V(d,d)$

Proof. Please see Bierman (1977) for instance. □

Note : An upper triangular factorisation follows a similar algorithm, changing only indices order.

C.1.2 The Jacob Method

Definition C.2. A $d \times d$ matrix V is said to have an eigen-vector $\mathbf{x}$ and corresponding eigenvalue λ if $V\mathbf{x} = \lambda\mathbf{x}$. Multiplies of $\mathbf{x}$ are not considered distinct eigen-vectors and the zero vector is not considered to be an eigen-vector at all.

Properties :

(i) Existence : The associated characteristic equation $|V - \lambda I| = 0$, derived from the definition, is a d^{th} degree polynomial in λ whose roots are the eigenvalues, that proves that there exist always d (not necessarily distinct) eigenvalues. Consequently, by the definition, there are d corresponding eigen-vectors (not necessarily distinct).

(ii) If V is symmetric, the eigenvalues of V will be real. Also, if they are distinct, the eigen-vectors $\mathbf{x}$ will be orthogonal to each other, forming an orthogonal matrix X. Then, by definition, $VX = X\ \mathrm{diag}(\lambda_1, \ldots \lambda_d)$, or equivalently, $X^{-1}VX = \mathrm{diag}(\lambda_1 \ldots \lambda_d)$, which is a special (similarity) transformation called diagonalization.

A more detailed description of this technique and full references can be found in Press (1986).

Appendix D

Notation and abbreviations

A summary of the most frequently used notation and abbreviations appears in the tables below. I also adopt the standard convention that random variables are denoted as capital letters, and instantiations of random variables (values) are denoted as lower-case letters. I do not explicitly distinguish between scalars and vectors. The vectors are denoted by bold-case letter and the scalers are without bold-case but the same letter as the corresponding vector. Throughout the thesis, I denote vectors (which may be random variables) as bold-case letter, to distinguish them from matrices, which are always italics.

Table D.1: Notation for general HBST model.

Symbol	Meaning
$\mathbf{Y}_t$	Observational vector at time t
$\mathbf{\Theta}_t$	State vector at time t
n	Total number of grid points
m	Total number of measuring stations
$\mathbf{Y}_{Dt}$	Observational vector at grid points
$\mathbf{Y}_{Mt}$	Observational vector at measuring stations
$\mathbf{\Theta}_{Dt}$	State vector associated with $\mathbf{Y}_{Dt}$
$\mathbf{\Theta}_{Mt}$	State vector associated with $\mathbf{Y}_{Mt}$
G_t	Spatial interpolation matrix
H_t	Time evolution matrix
$\{\nu_t\}$	Sequence of observational error vectors
$\{\epsilon_t\}$	Sequence of spatial interpolation error vectors
$\{\omega_t\}$	Sequence of time evolution error vectors
V_t	Observational variance-covariance matrix
Σ_t	Spatial interpolation variance-covariance matrix
W_t	Time evolution variance-covariance matrix
$\mathbf{I}_t = \{\mathbf{Y}_t, \mathbf{I}_{t-1}\}$	Information (observations) available up to time t
d_{ij}	Distance between sites i and j
d^*_{max}	Maximum distance from measuring stations to their nearest grid points
d_{max}	Maximum distance between grid points
Ψ_t	Scale matrix for inverse-Wishart distribution
$E(X)$	Expectation of a random variable X
$Var(X)$	Variance of a random variable X
$Cov(X1, X2)$	Covariance between two random variable $X1$ & $X2$
$X \amalg Y$	Random variables X and Y are independent
$X \amalg Y \mid Z$	Random variables X and Y are conditionally independent on Z

Table D.2: List of abbreviations.

Abbreviation	Meaning
ADM	Atmospheric Dispersal Model
Atm. Disp.	Atmospheric Dispersal
Cs134	Caesium-134
Cs137	Caesium-137
DLM	Dynamic Linear Model
Equip. data	Equipment data
Exp. Judg.	Expert Judgement
Food Chain Cont.	Food Chain Contamination
γ-ray data	Gamma-ray data
GCL	Ground Contamination Level
GIS	Geographic Information System
GMRF	Gaussian Markov Random Field
Ground Depos.	Ground Deposition
HBST Model	Hierarchical Bayesian Space-Time Model
Human Expos.	Human Exposures
I131	Iodine-131
IW	Inverse-Wishart
NIW	Normal Inverse-Wishart
MDLM	Multivariate Dynamic Linear Model
Ru103	Ruthenium-103
Ru106	Ruthenium-106
Te132	Tellurium-132

Bibliography

Anderson, T. W. (1984). *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons: New York, 2nd edition.

ApSimon, H., Goddard, A. J. H., and Wrigley, J. (1985a). Long-range atmospheric dispersion of radioisotopes-i. the mesos model. *Atmospheric Environment*, **19**(1), 99–111.

ApSimon, H., Goddard, A. J. H., Wrigley, J., and Crompton, S. (1985b). Long-range atmospheric dispersion of radioisotopes-ii. the mesos model. *Atmospheric Environment*, **19**(1), 113–125.

ApSimon, H., Wilson, J., and Simms, K. (1989). Analysis of the dispersal and deposition of radionuclides from chernobyl across europe. *Proc. R. Soc. Lond. A*, **425**, 365–405.

Atkinson, A. (1969). The use of residuals as a concomitant variable. *Biometrika*, **56**, 33–41.

Barbosa, E. and Harrison, J. (1992). Variance estimation for multivariate dynamic linear models. *Journal of Forecasting*, **11**, 621–628.

Barbosa, E. P. (1989). Dynamic bayesian models for vector time series analysis & forecasting. Unpublished PhD Thesis, University of Warwick.

Bartlett, M. S. (1938). The approximate recovery of information from replicated field experiments with large blocks. *Journal of Agricultural Science (Cambridge)*, **28**, 418–427.

Bartlett, M. S. (1978). Nearest neighbour models in the analysis of field experiments. *Journal of the Royal Statistical Society B*, **40**, 147–158.

Berger, J. O., Oliveira, V. D., and Sansó, B. (2001). Objective bayesian analysis of spatially correlated data. *Journal of the American Statistical Association*, **96**(456), 1361–1374.

Berliner, L. M. (1996). Hierarchical bayesian time series models. Maximum Entropy and Bayesian Methods, K. Hanson and R. Silver, Eds., Kluwer Academic Publishers, 15-22.

Besag, J. E. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B*, **36**, 192–225.

Besag, J. E. and Kempton, R. A. (1986). Statistical analysis of field experiments using neighboring plots. *Biometrics*, **42**, 231–251.

Bierman, G. J. (1977). *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York.

Bleher, M. and Jacob, P. (1993). Assessment of radionuclide deposition by park. *Radiation Protection Dosimetry*, **50**, 343–348.

Böllmann, U., Eder, E., Starke, H., Stein, H., and Zeising, H. (1987). Auswirkungen des reaktorunfalls in tschernobyl auf bayern. schriftenreihe. Heft **82**. Bayerisches Landesamt fur Umweltschutz.

Box, G., Jenkins, G., and Reinsel, G. (1994). *Time Series Analysis: Forecasting and Control*. Englewood Cliffs: Prentice Hall, 3rd edition edition.

Carroll, R., Chen, R., George, E., Li, T., Newton, H., Schmiediche, H., and Wang, N. (1997). Ozone exposure and population density in harris county. *J. Am. Statist. Ass.*, **92**, 392–415.

Castronovo, F. P. (1999). Teratogen update: Radiation and chernobyl. *Teratology*, **60**, 100–106.

Clayton, D. and Kaldor, J. (1987). Empirical bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics*, **43**, 671–681.

Cliff, A. D. and Ord, J. K. (1981). *Spatial Processes: Models and Applications*. Pion Limited.

Cook, D. G. and Pocock, S. J. (1983). Multiple regression in geographic mortality studies, with allowance for spatially correlated errors. *Biometrics*, **39**, 361–371.

Cressie, N. A. C. (1993). *Statistics for Spatial Data*. John Wiley and Sons, NY, revised edition edition.

Dawid, A. (1979). Conditional independence in statistical theory. *J. R. Statist. Soc. B*, **41**(1), 1–31.

De Oliveira, V., Kedem, B., and Short, D. A. (1997). Bayesian prediction of transformed gaussian random fields. *Journal of the American Statistical Association*, **92**, 1422–1433.

Dellaportas, P., Denison, D., and Mallick, B. (1998). Space-time modelling without distance. Technical report, Department of Statistics, Texas A & M University, College Station.

Desiato, F. (1992). A long-range dispersion model evaluation study with chernobyl data. *Atmospheric Environment*, **26A**(15), 2805–2820.

Devell, L., Güntay, S., and Powers, D. A. (1995). The Chernobyl reactor accident source term: Development of a consensus view. De95, CSNI Report in preparation, OECD/NEA, Paris.

Fairfield, S. (1938). An empirical law describing heterogeneity in the yields of agricultural crops. *Journal of Agricultural Science (Cambridge)*, **28**, 1–23.

Fisher, R. (1935). *The Design of Experiments*. Oliver and Boyd, Edinburgh.

Gamerman, D., Salazar, E., and Reis, E. (2006). Some applications of dynamic gaussian process priors to the analysis of space-time data models. Valencia. ISBA 8th World Meeting on Bayesian Statistics, Benidorm (Alicante, Spain).

Gelfand, A., Ghosh, S., Knight, J., and Sirmans, C. (1998). Spatio-temporal modeling of residential sales data. *J. Bus. Econ. Statist*, **16**, 312–321.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (1995). *Bayesian Data Analysis*. Chapman and Hall.

Goldstein, M. (1976). Bayesian analysis of regression problems. *Biometrika*, **63**, 51–58.

Green, P. J., Jennison, C., and Seheult, A. H. (1985). Analysis of field experiments by least squares smoothing. *Journal of the Royal Statistical Society B*, **47**, 299–315.

Guttorp, P., Meiring, W., and Sampson, C. (1994). A space-time analysis of ground-level ozone data. *Environmetrics*, **5**, 241–254.

Handcock, M. and Wallis, J. (1994). An approach to statistical spatio-temporal modeling of meteorological fields. *J. Am. Statist. Ass.*, **89**, 368–390.

Handcock, M. S. and Stein, M. L. (1993). A bayesian analysis of kriging. *Technometrics*, **35**, 403–410.

Harrison, P. J. and Stevens, C. (1976). Bayesian forecasting (with discussion). *J. Roy. Statist. Soc. B*, **38**, 205–247.

Haslett, J. and Raftery, A. E. (1989). Space-time modelling with long-memory dependence: Assessing ireland's wind power resource (with discussion). *Appl. Statist.*, **38**, 1–50.

Higdon, D. (1999). A process-convolution approach to modeling temperatures in the north atlantic ocean. *Environ. Ecol. Statist.*, **5**, 173–190.

Hill, J. R., Hinkley, D. V., Kostal, H., and Morris, C. N. (1984). Spatial estimation from remotely sensed data via empirical bayes models. In Proceedings of the NASA Symposium on Mathematical Pattern Recognition and Image Analysis, L. F. Guseman, Jr., ed. Department of Mathematics, Texas A & M, 115-136.

Hirst, D., Storvik, G., and Syversveen, A. R. (2003). A hierarchical modelling approach to combining environmental data at different scales. *Appl. Statist*, **52**(3), 377–390.

Huang, H. C. and Cressie, N. (1996). Spato-temporal prediction of snow water equivalent using the kalman filter. *Comput. Statist. Data Anal.*, **22**, 159–175.

James, W. and Stein, C. (1960). Estimation with quadratic loss. In Proceedings of the Fourth Berkeley Symposium 1, ed. J. Neyman, 361-380. Berkeley: University of California Press.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, Series D*, **82**, 34–45.

Lauritzen, B. and Mikkelsen, T. (1999). A probabilistic dispersion model applied to the long-range transport of radionuclides from the chernobyl accident. *Atmospheric Environment*, **33**, 3271–3279.

Lauritzen, B., Baklanov, A., Mahura, A., Mikkelsen, T., and Sorensen, J. H. (2006). K-model description of probabilistic long-range atmospheric transport in the northern hemisphere. *Atmospheric Environment*, **40**, 4352–4369.

Lavine, M. (1998). Another look at conditionally gaussian markov random fields. *Bayesian Statistics*, **6**, 371–387.

Lavine, M. and Lozier, S. (1999). A markov random field spatio-temporal analysis of ocean temperature. *Environ. Ecol. Statist.*, **6**, 249–273.

Le, N. D. and Zidek, J. V. (1992). Interpolation with uncertain spatial covariances : A bayesian alternative to kriging. *Journal of Multivariate Analysis*, **43**(2), 351–374.

Mardia, D., Goodall, C., Redfern, E., and Alonso, F. (1998). The kriged kalman filter. *Test*, **7**, 217–285.

Masreliez, C. J. (1975). Approximate non-gaussian filtering. *IEEE Trans. Automat. Control*, **20**, 107–110.

Matérn, B. (1986). *Spatial Variation*. Berlin: Springer-Verlag, 2nd edition.

Mikkelsen, T., Larsen, S. E., and Thykier-Nielsen, S. (1984). Description of the risø puff diffusion model. *Nuclear Safety*, **67**, 56–65.

Müller, H. and Pröhl, G. (1993). Ecosys-87: A dynamic model for assessing radiological consequences of nuclear accidents. *Health Physics*, **64**, 232–252.

NEA (2002). Nuclear energy agency. Chernobyl: Assessment of radiological and health impacts. (Available from www.nea.fr/html/rp/chernobyl/chernobyl.html.).

Papadakis, J. (1937). Methode statistique pour des experiences sur champ. *Institut d'Amelioration des Plantes a Thessaloniki (Grece), Bulletin Scientifique*, **23**.

Päsler-Sauer, J. (1986). Comparative calculations and validation studies with atmospheric dispersion models. KIK Research Report 4146, Kernforchungszentium, Karlsruhe.

Päsler-Sauer, J. (1997). Description of the atmospheric dispersion model atstep. Technical report, RODOS(WG2)-TN(97)-01, Forschungszentrum Karlsruhe.

Pfeifer, P. and Deutsch, S. (1980a). Independence and sphericity tests for the residuals of space-time arma models. *Communs Statist. Simuln. Computn.*, **9**, 533–549.

Pfeifer, P. and Deutsch, S. (1980b). Stationarity and invertibility regions for low order starma models. *Communs Statist. Simuln. Computn.*, **9**, 551–562.

Picard, R. R. and Cook, R. D. (1984). Cross-validation of regression models. *Journal of the American Statistical Association*, **79**(387), 575–583.

Politis, K. and Robertson, L. (2004). Bayesian updating of atmospheric dispersion after a nuclear accident. *Appl. Statist.*, **53**(4), 583–600.

Press, W. H. (1986). *Numerical Recipes*. Cambridge University Press, Cambridge.

Quintana, J. M. (1985). A dynamic linear matrix-variate regression model. Technical Report 83, Research Report, Department of Statistics, University of Warwick, UK.

Quintana, J. M. and West, M. (1987). An analysis of international exchange rates using multivariate dlm's. *The Statistician*, **36**, 275–281.

Sampson, P. D. and Guttorp, P. (1992). Nonparametric estimation on nonstationary spatial covarince structure. *J. Am. Statist. Ass.*, **87**, 108–119.

Sansó, B. and Guenni, L. (1999). Venezuelan rainfall data analysed by using a bayesian space-time model. *Appl. Statist.*, **48**, 345–362.

Schlather, M. (1999). Introduction to positive definite function and to unconditional simulation of random fields. Technical report, Dept. of Mathematics and Statistics, Lancaster University, UK.

Shaddick, G. and Wakefield, J. (2002). Modelling daily multivariate pollutant data at multiple sites. *Appl. Statist.*, **51**(3), 351–372.

Smedley, C., Grindon, E., Dutton, L., and Vleeshhouwer, D. (1996). Source term estimation based on plant status. RODOS(WG5)-TN(**96**)-02.

Smith, J. (1989). Influence diagrams for statistical modelling. *The Annals of Statistics*, **17**(1), 654–672.

Smith, J. and French, S. (1993). Bayesian updating of atmospheric dispersion models for use after an accidental release of radioactivity. *The Statistician*, **42**, 501–511.

Smith, R. and Robinson, P. (1997). A bayesian approach to the modelling of spatio-temporal precipitation data. *Lect. Notes. Statist.*, **121**, 237–269.

Stein, C. (1955). Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In Proceedings of the Third Berkeley Symposium 1, ed. J. Neyman, 197-206. Berkeley: University of California Press.

Stein, M. L. (1999). *Interpolation of Spatial Data*. New York : Springer.

Stoffer, D. (1986). Estimation and identification of space-time armax models in the presence of missing data. *J. Am. Statist. Ass.*, **81**, 762–772.

Stroud, J. R., Müller, P., and Sansó, B. (2001). Dynamic models for spatiotemporal data. *J. R. Statist. Soc. B*, **63**(4), 673–689.

Student (1907). On the error of counting with a haemacytometer. *Biometrika*, **5**, 351–360.

Thykier-Nielsen, S., Deme, S., and Mikkelsen, T. (1998). RIMPUFF user guide: version RIMDOS8 (Risø National Laboratory, Roskilde, Denmark).

Tonellato, S. (1997). Bayesian dynamic linear models for spatial time series. Technical Report 5, Dipartimento di Statistica, Università Ca' Foscari di Venezia, Venice.

Tonellato, S. (1998). Spatial prediction with space-time models. Technical Report 2, Dipartimento di Statistica, Università Ca' Foscari di Venezia, Venice.

tutiempo.net (2008). (http://www.tutiempo.net/en/Weather/Germany/DE.html).

Upton, G. J. G. and Fingleton, B. (1995). *Spatial Data Analysis by Example vol.1: Point Pattern and Quantitative Data*. Wiley, reprint.

Wackernagel, H. (1995). *Multivariate Geostatistics*. Berlin: Springer-Verlag.

Walker, S. and Muliere, P. (1997). Beta-stacy processes and a generalization of the polya-urn scheme. *The Annals of Statistics*, **25**(4), 1762–1780.

Waller, L., Carlin, B., Xia, H., and Gelfand, A. (1997). Hierarchical spatio-temporal mapping of disease rates. *J. Am. Statist. Ass.*, **92**, 607–617.

Watari, K., Imai, K., Ohmomo, Y., Muramatsu, Y., Nishimura, Y., Izawa, M., and Baciles, L. R. (1988). Simultaneous adsorption of cs-137 and i-131 from water and milk on metal ferrocyanide-anion exchange resin. *Journal of Nuclear Science and Technology*, **25**(5), 495–499.

West, M. (1982). Aspects of recursive bayesian estimation. *Unpublished PhD thesis, University of Nottingham.*

West, M. and Harrison, P. J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York, 2nd edition.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, **41**, 434–449.

Wikle, C. and Cressie, N. (1999). A dimension reduced approach to space-time kalman filtering. *Biometrika*, **86**(4), 815–829.

Wikle, C., Berliner, M., and Cressie, N. (1998). Hierarchical bayesian space-time models. *Environ. Ecol. Statist.*, **5**(2), 117–154.

Wilkinson, G. N., Eckert, S. R., Hancock, T. W., and Mayo, O. (1983). Nearest neighbor (nn) analysis with field experiments. *Journal of the Royal Statistical Society B*, **45**, 151–178.

WNA (2008). World nuclear association: Chernobyl accident. (Available from http://www.world-nuclear.org/info/chernobyl/inf07.html.).

Yaglom, A. (1987). *Correlation Theory of Stationary and Related Random Functions I, Basic Results*. Springer-Verlag, New York, Berlin, Heidelberg.

Yates, F. (1938). The comparative advantages of systematic and randomized arrangements in the design of agricultural and biological experiments. *Biometrika*, **30**, 444–466.

Zimmerman, D. L. and Harville, D. A. (1989). On the unbiasedness of the papadakis estimator and other nonlinear estimators of treatment contrasts in field-plot experiments. *Biometrika*, **76**, 253–259.